Gabriel M. Ahlfeldt, Kristoffer Moeller and Nicolai Wendland

Chicken or egg? the PVAR econometrics of transportation

Article (Accepted version) (Refereed)

Gabriel M. Ahlfeldt[♣], Kristoffer Moeller[⊗] & Nicolai Wendland[♦]

Chicken or egg? The PVAR econometrics of transportation[*]

Abstract: To analyse the mutually dependent relationship between local economic performance, demand for and supply of transport services, we employ the structural panel VAR method that is popular in the macroeconomic literature, but has not previously been applied to the modelling of the within-city dynamics of transportation. We focus on a within-city panel of Berlin, Germany during the heyday of the construction of its dense public transit network (1890–1914). Our results suggest that economic outcomes and a supply of transport infrastructure mutually determine each other. We find a short-run (long-run) elasticity of property prices with respect to transport supply of 2% (8.5%). Both transport demand and supply seem to be driven more by firms than by residents.
Keywords: *transport, land use, Berlin, history, panel vector autoregression*
Version: *July 2014*
JEL: R12, R14, R41, N73, N74

1 Introduction

The relationship between transport infrastructure and various economic outcomes is plagued by a notorious simultaneity problem. Put simply, there are plenty of reasons to believe that an ease of access to other locations within a city, region, country, or beyond should have a positive impact on the attractiveness of a location. At the same time the demand required to recover large investments into infrastructures like airports, highways or railways is most likely strongest among economically successful places. The allocation of transport infrastructure is therefore non-random in most cases, which has long complicat-

♣ *Corresponding author:* London School of Economics and Political Sciences & Spatial Economics Research Centre (SERC). g.ahlfeldt@lse.ac.uk, www.ahlfeldt.com

⊗ Darmstadt University of Technology and Center for Metropolitan Studies TU Berlin mail@kristoffer-moeller.de, www.kristoffer-moeller.de.

♦ Darmstadt University of Technology. wendland@vwl.tu-darmstadt.de

* We thank Bill Kerr (the editor), one anonymous referee and seminar and conference participants in London (SERC), Bratislava/Palermo (ERSA) and Duisburg (EGIT) and especially Kurt Schmidheiny, Henry Overman, Jens Südekum, Volker Nitsch and Sevrin Waights for helpful comments and suggestions. Sascha Möbius and Neele Reimann-Phillip also provided excellent research assistance. The German Science Foundation (DFG project code NI 938/2-1) and the Fritz-Thyssen-Foundation (project code AZ.10.10.2.070) are acknowledged for financial support. The usual disclaimers apply.

ed the interpretation of the evident positive correlation between accessibility and economic performance. A priori, it is not clear to which extent this correlation is attributable to economic impact caused by the supply of transport infrastructure, or by transport supply being provided in response to demand, a bidirectionality that resembles the chicken-or-egg problem. Moving beyond correlation and towards establishing causality, however, it is important to justify (public) expenditures on transport infrastructure that are supposed to promote local economic development.

Our approach to deal with the simultaneity of economic outcomes, demand for and supply of transportation is borrowed from a macro-econometric literature that faces similar econometric challenges, albeit in different economic contexts. In this literature, structural vector autoregressive (VAR) models are often used to model the complex temporal structure among economic outcome variables and policy variables that typically reflect causes and effects of economic performance. Typical fields of application of the panel version of VAR (PVAR) include monetary policy and investment behavior (Assenmacher-Wesche & Gerlach, 2008; Carlino & DeFina, 1998; Love & Zicchino, 2006), supply of development aid (Gillanders, 2011; Gravier-Rymaszewska, 2012; M'Amanja & Morrissey; Osei, Morrissey, & Lloyd, 2005) or security economics (Konstantinos & Konstantinou, 2013). Some applications of the VAR method have focussed on regional transport analyses, especially with respect to the crowding-in effects of public/private investment (Bogart, 2009; Pereira & Andraz, 2012a, 2012b; Pereira & de Frutos, 1999). In a nutshell, the (panel) VAR method is particularly useful if the theory offers some guidance as to the potential directions and the temporal sequence of the dependencies among variables, but does not deliver predictions that translate into a unidirectional impact of (an) exogenous variable(s) on a dependent variable. Identification is achieved by imposing restrictions on the temporal structure of causalities, i.e. by ruling out contemporary shocks of demand on supply.

With this contribution, we make the case that the PVAR method can be applied to the chicken-or-egg problem of transportation, which resembles typical macro-policy problems at a spatial micro-level. We apply the PVAR method to a panel data set containing measures of land value, population density and transport services that is rich in spatial detail and spatiotemporal variation. We argue that assuming a concave production function in the construction sector there is a bidirectionally unambiguous relationship between the supply of transport services and the value of urban land. The chicken-or-egg problem in transport economics can therefore in principle be investigated based on an analysis of these two variables alone. We further argue that (assuming land markets clear) a data set containing land value, population and transport supply measures can be used to analyse the extent to which firms or residents respond to transport shocks and whether transport infrastructure is in practice directed towards firms or residents.

Rich variation in transport supply is typically difficult to find for contemporary cities of the developed world. Baum-Snow et al. (2012) have responded to that limitation by analysing economic growth and transport development patterns in a developing country (China) using a long-difference analysis. Limited data availability, however, makes it difficult to build a spatial panel with sufficient temporal coverage and spatial detail for a PVAR analysis of a developing country. Our analysis, therefore, remains located in the developed world, but is set in a period where many of today's established cities were themselves developing. The focus of our analysis is on Berlin, Germany, 1890–1914, which was not only a period of massive economic growth—the population about doubled during this period—but was also the period when the backbone of today's within-city transit network was developed. A further important advantage of the historic setting is that automobiles can be ignored as a relevant transport mode. Our analysis benefits from one of the (world-wide) rare data sources offering land values at a high spatial detail and for various points in time: the so-called Müller maps, which present detailed categories of assessed land values on highly disaggregated geographical levels up to individual plots.[1] While selected Müller land values have been utilised in previous research (Ahlfeldt & Wendland, 2009, 2011, 2013) this is the first application of a complete digitised record of the information contained in these maps. We complement the land value information with a neighbourhood panel of population records and information on historic transport networks that were digitally reproduced in GIS based on historic transport maps.

Besides the methodological similarity to the macro-econometric literature, our research directly connects to a wide range of urban economics research that aims at establishing the unidirectional causal impact of transport supply on economic outcome measures using either quasi-experimental (Ahlfeldt, 2013; Ahlfeldt & Wendland, 2009; Gibbons & Machin,

[1] Chicago is the notable exception in terms of availability of land value estimates. Hoyt and Olcott have provided spatially varying land value estimates covering a period of about 150 years (see McMillen, 1996 for details).

2005; McDonald & Osuji, 1995; McMillen & McDonald, 2004; Michaels, 2008) or instrumental variable (IV) designs (Baum-Snow, 2007; Baum-Snow, et al., 2012; Duranton & Turner, 2011, 2012; Holl & Viladecans-Marsal, 2011; Hornung, 2012; Hsu & Zhang, 2011). These studies typically implicitly or explicitly assume that the supply of infrastructure is uncorrelated with the previous trend in observed economic outcome in a particular case or that an IV is at hand that predicts transport supply, but is conditionally uncorrelated with the outcome. Provided the identifying assumptions are met, these approaches typically identify a positive causal effect of transport supply on an economic outcome at the expense of not being informative with respect to the feedback of economic development on transport supply. Fewer studies have provided evidence of the impact of economic development on the provision of transport infrastructure (Cervero & Hansen, 2002; Levinson & Karamalaputi, 2003).

In the sense that our method explores the bidirectional temporal relationship between transport demand and supply variables, our approach is most closely related to Levinson's (2008), Xie & Levinson's (2010) and Granger's (1969) causality analyses of the relationship between transport services supply and population density. Their results point to co-development, i.e. a mutually causal relationship between demand for and supply of transport. The key advantage of our data set relative to their analysis is the additional availability of land value data. As we discuss in more detail in the next section, the (mutual) relationship between transport supply and land value is presumably unambiguously positive, whereas the relationship between transport supply and residential use in both directions is affected by the competing commercial use. Combining measures of land value, transport supply and residential density, moreover, allows for additional insights to be gained into whether the relationship between transport demand and supply is driven primarily by residential or commercial use. Compared to a Granger causality test, the PVAR method we employ has the advantage of being able to accommodate multi-directional causal impact in multivariate systems of equations and allows for a structural interpretation of reduced form coefficients under the assumption that the temporal structure of mutual dependencies of the endogenous variables is known. Methodologically, our approach is closely related to Graham et al. (2010) who make use of a PVAR model to test the direction of (Granger) causality between agglomeration economies and productivity. In (Melo, Graham, & Canavan, 2012) a PVAR approach is applied to examine the link between investment in road transport and economic output allowing for simultaneously induced travel demand.[2]

Previewing our findings, our application of the PVAR methods suggests that the relationship between economic performance and transport supply within our research environment is bidirectional, that commercial activity tends to displace residential use in response to transport improvements, and that transport planners have followed commercial activity more than residential demand, especially in the economic core of our study area.

The remainder of the paper is structured as follows. In the next section we provide some theoretical guidance to the interpretation of our PVAR model results and show how we take the model to the data. Section 3 presents the results and the final section concludes.

2 Empirical strategy

To analyse the relationship between transport demand and supply researchers have relied on different indicators of economic activity. Some researchers have inferred transport demand from house or land price capitalisation (e. g. Ahlfeldt & Wendland, 2011; Gibbons & Machin, 2005). Other researchers have focussed on population (density) instead (e.g. Levinson, 2008). The purpose of the next subsection is to provide a simple theoretical framework that helps with the interpretation of bivariate and multivariate PVAR models which use different indicators of local economic demand. We will discuss three different settings. Firstly, a bivariate model using land values and a measure of transport supply that is sufficient to model the dynamics of demand and supply. Secondly, an alternative bivariate model which uses population density instead of land values. In this model the analysis then needs to be carried out separately by land use. The reason is that depending on the relative attractiveness of an area to firms and residents, the population will either be attracted to an area with improved transport supply, or outbid and thus displaced by firms that relocate to the area. The third setting involves a multivariate model which includes land values, population density and a measure of transport supply. This model is

---

[2] Among the few applications of the PVARs method in urban/regional economics are Miller & Peng (2006), Lee (2007) and (Brady, 2011). (Calderón, Moral-Benito, & Servén, 2014) analyse the relationship between economic outcome and infrastructure capital using panel time-series techniques.

the most difficult to interpret since the effects of the two demand-side variables need to be interpreted conditional on each other. The advantage of this model is that it allows for some insights into the causes and effects of changes in the land use pattern that remain hidden in a two variable land value transport supply model.

A motivating theoretical framework, the PVAR method and how we take it to our data are discussed in sections 2.1, 2.2, and 2.3. Section 2.4 then explains the nature of the historic data that we collected.

2.1 Theoretical framework

In our stylised world, accessibility positively shifts the local residential and commercial demand for space. The supply curve of usable floor space is upward sloping because the supply of land at a given location is fixed and the available construction technology imposes limits to densification. Accessibility therefore increases the equilibrium price and quantity of space consumed as well as the land value at a given location. We assume that there is a unique mapping of equilibrium land value to equilibrium space (per land unit). As residential and commercial use are mutually exclusive the land use mix is therefore exactly identified by a measure of land value and a measure of either the commercial or residential use.

Let's assume that local demand for residential space q^P and commercial space per land unit q^E is defined by the following demand functions:

$$q^P = q^P(v,T,L), \frac{\partial q^P}{\partial v} < 0, \frac{\partial q^P}{\partial T} > 0, \frac{\partial q^P}{\partial L} > 0 \quad (1)$$

$$q^E = q^E(v,T,L), \frac{\partial q^E}{\partial v} < 0, \frac{\partial q^E}{\partial T} > 0, \frac{\partial q^E}{\partial L} > 0 \quad (2),$$

where v is the price of a homogenous unit of building floor space, T is a measure of locally available transport services and L captures time-invariant locational features that make a location more attractive. Demand for space is decreasing in the price of space and increasing in the quality of transport services and amenities. All variables are expressed in per land unit terms.

We further assume a competitive construction sector with a concave production function. As demonstrated by Epple et al. (2010), the price of land – the land value V – must be a monotonic function of the price of a homogenous unit of building space as long as the unit price of developed land is a monotonic function of the price of building space, i.e. v=v(V). The local supply of building space q^S increases at a decreasing rate in land value reflecting increasing incentives to use non-land inputs and limits to densification.

$$q^S = q^S(V), \frac{\partial q^S}{\partial V} > 0, \frac{\partial^2 q^S}{\partial V^2} < 0 \tag{3}$$

Market clearing implies that total demand for residential and commercial space must equal supply.

$$q^P + q^E = q^S \tag{4}$$

The market clearing condition allows us to derive the equilibrium quantities Q^{E^*} and Q^{P^*}, which add up to Q^{S^*}, and the equilibrium land value V^*. Assuming that the consumption of space per firm and household is constant Q^{E^*} can be approximated by the local number of firms (alternatively as local employment) and Q^{P^*} by the local number of residents (population density), respectively.

Transport planners design the infrastructure to accommodate demand such that transport services T increase in the number of local residents Q^{P^*} and firms Q^{E^*}.

$$T = T\left(Q^{P^*}, Q^{E^*}\right), \frac{\partial T}{\partial Q^{E^*}} > 0, \frac{\partial T}{\partial Q^{P^*}} > 0 \tag{5}$$

Assuming market clearing (4) and the monotonicity of the relationship of the price of usable space and land value (v=v(V)), equations (1–3) jointly determine the equilibrium land value as a function of transport access and location amenities ($V^* = V^*(T, L)$).[3] Since T is itself a function of Q^P and Q^E and market clearing implies that $Q^{E^*} = Q^{S^*} - Q^{P^*}$ we can express the land value as:

$$V^* = V^*\left(T\left(Q^{P^*}, Q^{S^*}(V^*) - Q^{P^*}\right), L\right)) \tag{6}$$

It is further straightforward to re-express transport access as a function of land value and the local number of residents Q^P:

$$T^* = T^*\left(Q^{P^*}, Q^{S^*}(V^*) - Q^{P^*}\right) \tag{7}$$

[3] See Brueckner (1987) for a detailed derivation of the urban equilibrium determined by the demand (Alonso, 1964; Fujita & Ogawa, 1982; Lucas & Rossi-Hansberg, 2002) and supply side (Epple, et al., 2010; Mills, 1972; Muth, 1969; Saiz, 2010) of land markets.

Since residential demand for space is a function of transport access, equilibrium land value and time-invariant amenities (1), it is possible to further simplify the equation system.

$$V^* = V\big(T^*(Q^{P^*}(V^*, T^*, L), Q^{S^*}(V^*) - Q^{P^*}(V^*, T^*, L)), L\big) = V^*(T^*, L),\ \frac{\partial V^*}{\partial T^*} > 0 \tag{8}$$

$$T^* = T^*\big(T^*(Q^{P^*}(V^*, T^*, L), Q^{S^*}(V^*) - Q^{P^*}(V^*, T^*, L)), L\big) = \mathrm{T}^*(V^*, L), \frac{\partial T^*}{\partial V^*} > 0 \tag{9}$$

A nice feature of the bidirectional relationship between land value and transport access is that the mutual influence is theoretically unambiguously positive. Note that the panel nature of our data set allows us to hold the effect of time-invariant amenities L constant so that we abstract from L in the remainder of this section.

We chose the historic environment for this study due to the substantial spatiotemporal variation in transport access compared to most contemporary cities with already developed mass transit. Unfortunately, the historic environment also implies that it is more difficult to collect data that are sufficiently spatiotemporally disaggregated to be suitable for the relatively demanding PVAR method. It has proven impossible to collect spatiotemporally disaggregated data on employment or building stock that would be suitable for the analysis. What we are able to observe besides land values and a self-constructed measure of transport access is a within-city panel of population density. This is already a relatively rich data set given the historical setting.

With a concave production function in the construction sector and the market clearing condition (4), introduced above, however, it is possible to rearrange equations (1–4) such that a unique equilibrium is defined with three (land values, transport access, population density) or even two (land values, transport access) variables (as shown above). Equation (6) can be rearranged to give:

$$V^* = V^*\big(Q^{P^*}, T^*, L\big), \frac{\partial V^*}{\partial T^*} > 0, \frac{\partial V^*}{\partial Q^{P^*}} \lesseqgtr 0 \tag{10}$$

A positive relationship is expected between the two endogenous variables V^* and T, but the expectations are unclear regarding the effect of population density on land value conditional on transport access. The intuition is that the relationship between population density and land value critically depends on the degree to which transport access attracts firms, relative to population $\left(\frac{\partial Q^{P^*}}{\partial T^*} \lesseqgtr \frac{\partial Q^{E^*}}{\partial T}\right)$.

Similarly, equation (7) can be rearranged to give:

$$T^* = T^*\left(Q^{P^*}, V^*\right), \frac{\partial T^*}{\partial V^*} > 0, \frac{\partial T^*}{\partial Q^{P^*}} \lesseqqgtr 0 \quad (11)$$

Since $\partial T^*/\partial Q^{P^*} > 0$ and $\partial Q^{S^*}/\partial V^* > 0$ we unambiguously expect the land value to be a positive determinant of transport access. Again, however, the expectation is ambiguous for the relationship with population density.

The lack of data on employment and housing stock also complicates the interpretation of the population response to changes in transport access and land value. Under market clearing assumptions, the space consumed by residents is determined by their valuation of transport access relative to firms' valuation.

$$Q^{P^*} = \begin{cases} Q^{P^*}(V^*, T^*) \\ Q^{S^*}(V^*) - Q^{E^*}(V^*, T^*) \end{cases} = Q^{P^*}(T^*, V^*), \frac{\partial Q^{P^*}}{\partial V^*} \lesseqqgtr 0, \frac{\partial Q^{P^*}}{\partial T^*} \lesseqqgtr 0 \quad (12)$$

There are generally no unambiguous theoretical expectations regarding the directions of the relationships involving population density. Nevertheless, equations (10–12) suggest that an empirical approximation facilitates interesting tentative interpretations. A positive (negative) impact of population density on land value caused by, e.g. a transport shock, will suggest a relatively higher (lower) WTP by residents vs. firms. Likewise, a positive (negative) impact of population density on transport access (conditional on land value) will be suggestive of a transport planner who targets residential (commercial) areas. Finally, positive (negative) population density responses to transport (and conditional land value) changes would again indicate a higher WTP by residents relative to firms.

From a theoretical point of view the interpretation of a bidirectional relationship between population density and transport access alone is in principle even more complicated than in the three variable system since the demand linkage of population and transport supply works in the opposite direction of the competition between population and (unobserved) employment. One simple way to address this problem, however, is to break the city down into areas where residents are expected to outbid firms (demand effect dominates) and vice versa (competition effect dominates) (Levinson, 2008).

2.2 Methodology

A VAR model consists of a system of equations which are estimated simultaneously. Each variable in this system is explained by its own lags and lagged values of the other variables.[4]

$$y_{i,t} = A_0 a_{i,t} + M_1 y_{i,t-1} + \cdots + M_p y_{i,t-p} + u_{i,t}\ , \tag{13}$$

$$(i = 1, \ldots, N; t = 1, \ldots . Z)$$

$$u_{i,t} = \mu_i + \nu_t + \varepsilon_{i,t} \tag{14}$$

where $y_{i,t}$ is a κ x 1 vector of κ panel data variables, the $M's$ are κ x κ coefficient matrices of the lagged variables $y_{i,t}$, p denotes the number of lags and $a_{i,t}$ is a vector of deterministic terms (linear trend, dummy or a constant) with the associated parameter matrix A_0. The unobserved individual effect μ_i, the time fixed effect ν_t and the disturbance term $\varepsilon_{i,t}$ jointly compose the error process $u_{i,t}$. We assume that $u_{i,t}$ has zero mean, i.e. $E(u_t) = 0$, independent u_t's and a time invariant covariance matrix.

We control for individual fixed effects by forward-mean-differencing (also Helmert transformation), i.e. we remove the mean of all future observations available for each location i –time t pair.[5] The Helmert transformation preserves the orthogonality between the variables and their lags which is essential for the use of lags as instruments in a system GMM estimation (Arellano & Bover, 1995).[6] We also time-demean all series to control for time effects.

Panel VAR estimation requires stationary variables. Acknowledging the short panel nature of our data set (N large, T small) we apply a modified Fisher-type unit root test. Following Choi (2001) we use the modified version of the inverse X^2 transformation in order to test the null hypothesis of all panels having a unit root.

4 See Sims (1980) and Holtz-Eakin (1988) for the theoretical underpinnings.

5 Applying standard mean-differencing procedures generates biased estimates as the fixed effects are correlated with the regressors due to the auto-correlated dependent variables (Arellano & Bond, 1991; Arellano & Bover, 1995; Blundell & Bond, 1998).

6 We use the STATA routines `pvar` and `helm` developed by Inessa Love for an econometric analysis by Love & Zicchino (2006). The original programmes are available at http://go.worldbank.org/E96NEWM7L0.

Based on the reduced form results and the moving average representation of the VAR model (Wold decomposition), the impulse response functions (IRF) can be derived to show how a variable reacts to a unit innovation in the disturbance term in period t holding all shocks constant. The confidence bands of the IRF are generated in Monte Carlo simulations following Love & Zicchino (2006).

2.3 Implementation

We estimate two alternative equation systems to empirically approximate the bi-lateral relationship of land value V and transport supply T as well as the multilateral relationship between land value V, transport supply T and population P. Since all variables are in logs and we identify from variation over time using constant geographies, a spatial normalisation (transformation into density) becomes empirically obsolete.

$$T_{i,t} = \theta_1 + \alpha_{11}V_{i,t-1} + \alpha_{12}T_{i,t-1} + v_{1i,t} \quad (15)$$

$$V_{i,t} = \theta_2 + \alpha_{21}V_{i,t-1} + \alpha_{23}T_{i,t-1} + v_{2i,t} \quad (16)$$

$$T_{i,t} = \delta_1 + \beta_{11}V_{i,t-1} + \beta_{12}P_{i,t-1} + \beta_{13}T_{i,t-1} + u_{1i,t} \quad (17)$$

$$P_{i,t} = \delta_2 + \beta_{21}LV_{i,t-1} + \beta_{22}P_{i,t-1} + \beta_{23}T_{i,t-1} + u_{2i,t} \quad (18)$$

$$V_{i,t} = \delta_3 + \beta_{31}V_{i,t-1} + \beta_{32}P_{i,t-1} + \beta_{33}T_{i,t-1} + u_{3i,t} \quad (19)$$

Since our short panel contains six time periods only, we set the lag length to $p = 1$. In the first step, the reduced form VAR systems described above are estimated using system GMM (Arellano & Bover, 1995). In the second step, we compute the IRF orthogonalising the residuals to move from the reduced form coefficients to a more structural interpretation. For the identifying restriction we impose the following recursive ordering of causality (see for details on the Choleski decomposition Enders, 1995; Hamilton, 1994): Transport (T), population (P), land values (V). The earlier a variable appears in the system the presumably more exogenous it is.

We suppose that the construction of a new transport infrastructure is not affected by any contemporaneous shocks, only by lagged variables. This is simply because it takes time to plan and build stations and networks ("time-to-build effects") (Kilian, 2011; Love & Zicchino, 2006) and we therefore don't expect instant responses to population or land value shocks. Population is assumed to react to contemporaneous transport shocks but not to contemporaneous shocks in land values. The construction of new lines is usually publicised in advance so that residents have time to adjust their location according to their preference for accessibility. Land value adjustments, in contrast, are typically not readily observed by residents. There are likely information delays (Inoue, Kilian, & Kiraz, 2009) which are assumed to last one period in our system. The supposedly least exogenous variable, land values, is assumed to be affected by contemporaneous transport and population shocks (plus lag of all variables). This is in line with the weak form efficient market hypothesis according to which markets incorporate all realisations of relevant outcomes (Fama, 1970). According to the semi-strong (or strong) efficient market hypothesis, however, markets are expected to immediately respond to any information made publicly available. This would imply an adjustment to the announcement and not the completion of a new rail station if – even in the early days of metro rail development – markets were able to forecast land price effects. We address this concern in a robustness check using artificial historic transport networks based on the announcement, not the inauguration, of new transport infrastructures.

We run and present the analysis in the following sequence. First, we concentrate on the bilateral relationship between a measure of transport supply introduced in more detail in the next subsection and land value, which we consider a global economic output measure in the sense that it reflects the productivity of land irrespectively of the type of use. At this stage we are able to make full use of the extraordinary spatial detail of the land value information at hand. Second, we explore the mutual relationships between land value, transport supply and population in a multivariate PVAR model. At this stage, we lose some of the spatial detail as we aggregate our data to the neighbourhood level for which population data is available. As discussed, we expect to gain insights into how different land uses (residential vs. commercial) respond to and impact on transport supply. Following the standard practice we report the reduced form estimates, the IRF and the variance decomposition for both PVAR models.

These main stages of the analysis are complemented by a range of robustness checks, model extensions and complementary analyses to cross-validate the implications drawn from the benchmark PVAR models. We repeat the benchmark estimates using transport variables generated based on announcement dates to assess the sensitivity of the results with respect to implicit market efficiency assumptions. We repeat the analysis using varying levels of spatial data aggregation and different measures of transport supply. We further break down the estimation samples into presumably commercial (downtown) and residential (periphery) areas where the theoretical implications regarding the population effects are less ambiguous. At this stage we explore the bivariate relationship between population and transport supply in more depth by making use of the longer time-dimension in these data series (compared to the land value data). Finally, the conclusions regarding the causal effects of transport supply on land value and land use are contrasted with the results of a complementary analysis using an instrumental variable approach.

2.4 Data

For the estimation of our PVAR models we make use of information on land value, population and transport infrastructure that are disaggregated by space and time. Our land value measures are constructed based on plot level land value maps published by Gustav Müller between 1890 and 1914. On these maps, various land value categories represented by items (e.g. circles, triangles, etc.) of different colours are assigned to individual plots that typically lie along a street front. The level of spatial detail is high. Categories typically change even within blocks of houses and across two sides of the same street and usually reflect the effect of being located at a corner. Each category corresponds to a numeric land value (e.g. 5 RM) that is specific to each issue. We digitise six cross-sections using the following procedure.[7] First, we scan and georeference the historic maps in GIS. Second, we draw lines (polylines) along road sections of the same land value category and assign the respective land value interval. Third, we aggregate the polyline values to spatial units (polygons for grid cells or neighbourhoods), weighted by the line length within a polygon.

All land values are given in Reichsmark per square meter. While Müller did not explicitly reveal the exact procedure of land value assessment, the imperial valuation law (Reichsbewertungsgesetz) of the German Reich contained strict orders to use capital values for the assessment of plots based on fair market prices. In line with the valuation laws, land values thus refer to the pure site value and are adjusted for all building and even garden characteristics that are not an intrinsic feature of the location. Müller also corrects for specific plot characteristics such as single and double corner lots, subsoil and courtyard properties. In many respects, the Müller data are comparable to Olcott's land values utilised by, for example, McMillen (1996). A subset of Müller values has been used by Ahlfeldt & Wendland (2009, 2011, 2013), who also provide a more detailed discussion of the data

[7] The land value data were extracted for 1890, 1896, 1900, 1904, 1910 and 1914.

along with some comparisons to later publications of land values in Berlin (Kalweit, 1928, 1936; Runge, 1950).

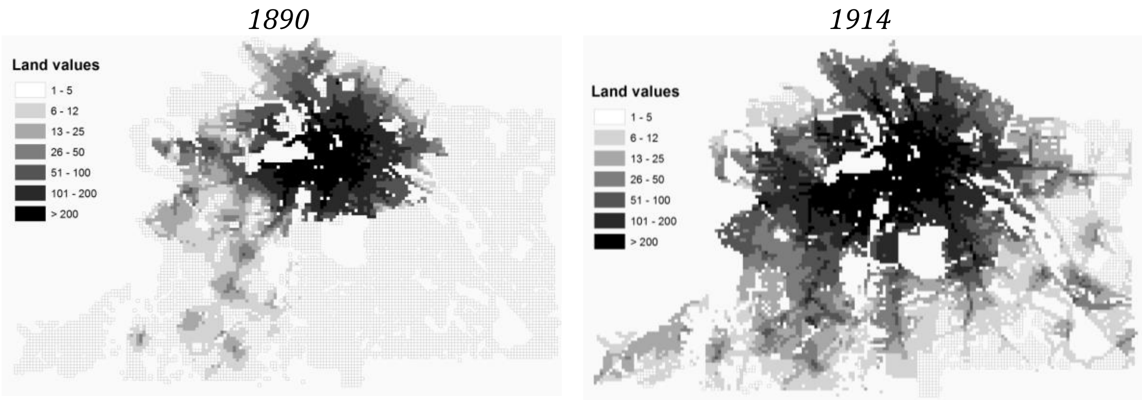
To make full use of the high spatial detail provided by the Müller maps, we aggregate the data to relatively fine 150x150m grid cells, which form the cross-sectional identifier in our bivariate PVAR models (see Figure 1). We note that due to the rapid expansion of the city during the study period the spatial extent covered the Müller map increases over time. To arrive at a balanced panel we assign the minimum land value observed in a given year to grid cells outside the city margin. This approach is rationalised by assuming that we observe some fraction of the urban margin where, in the absence of zoning regulation, the urban land value corresponds to the agricultural land rent (Alonso, 1964).

In comparison to the land value data, the population data we collect provides an even more comprehensive coverage (basically the whole of Berlin), but at a lower spatial detail (see Figure 2). We collect 14 cross-sections (in five year intervals from 1870 to 1935) of 93 neighbourhoods ("Ortsteile") from the Statistical Yearbook of Berlin (Statistisches Amt der Stadt Berlin, 1920).

Our measure of transport supply is an index of effective accessibility to stations connected to the heavy rail network in Berlin. This network effectively consist of two separate networks which are, however, close substitutes in terms of speed, comfort and transport fares. The suburban railway network (today the "S-Bahn") mostly connects central areas to suburban areas and locations along a semi-central circular line to each other. It was largely developed during the last decades of the 19th century. The underground ("U-Bahn") has formed a relatively dense network within central locations. Compared to the suburban rail network, it provides less coverage in the outer areas. Development of the U-Bahn did not start before 1902 (see also Table 1). We provide a detailed discussion of the history of the Berlin transport network in section 2 of the appendix.

The rationale for focussing on the two heavy rail systems is twofold. First, the heavy rail systems, with an average velocity of 33km/h (Ahlfeldt & Wendland, 2011), provided a significant accessibility advantage over any other transport mode. Second, a dense bus and streetcar network had already been developed by the end of the 19th century, implying that variations in accessibility over time were primarily driven by stations that were added to the heavy rail network. Similar to Levinson (2008) we capture the effective accessibility to the heavy rail network using a kernel density measure which discounts surrounding stations on distance.[8] Our kernel uses a radius of 2km (Silverman, 1986), which is in line with the catchment area of London underground stations identified by Gibbons and Machin (2005). Compared to a simple distance to the nearest station measure the kernel density measure incorporates the marginal benefit of having a second (or third, etc.) station in vicinity. To the extent that different stations are connected to different lines, they will also offer different and potentially complementary transport services. To compute our density measure for different years we have collected, scanned, and digitised historic networks in GIS using various historic sources (Mauruszat, 2010; Schomacker, 2009; Straschewski, 2011). Figure 3 illustrates the remarkable increase in effective access to the rail network over our observation period.

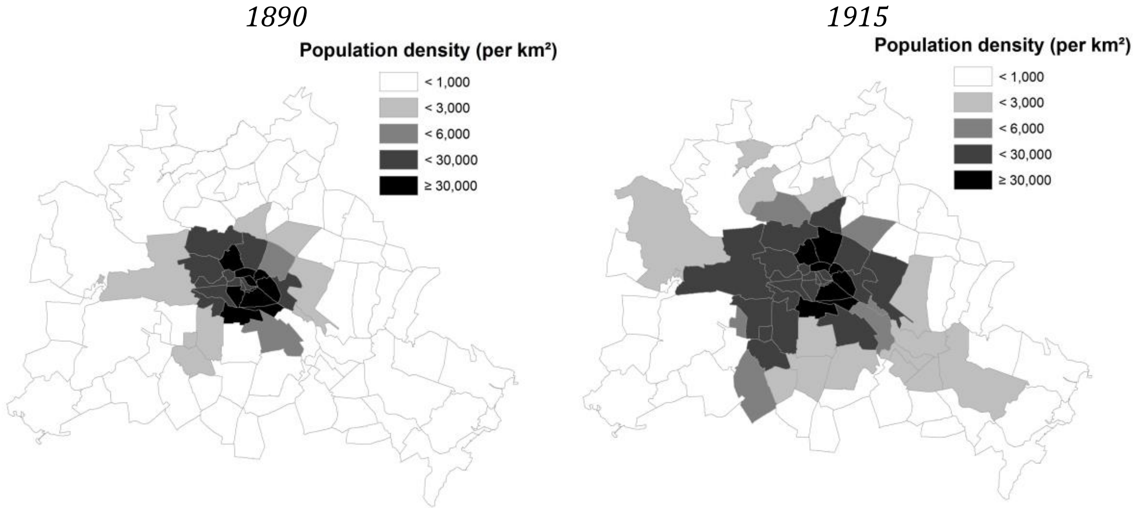
Fig. 1. Land values by grid cells



Note: Data are aggregated to 150x150m grid cells.

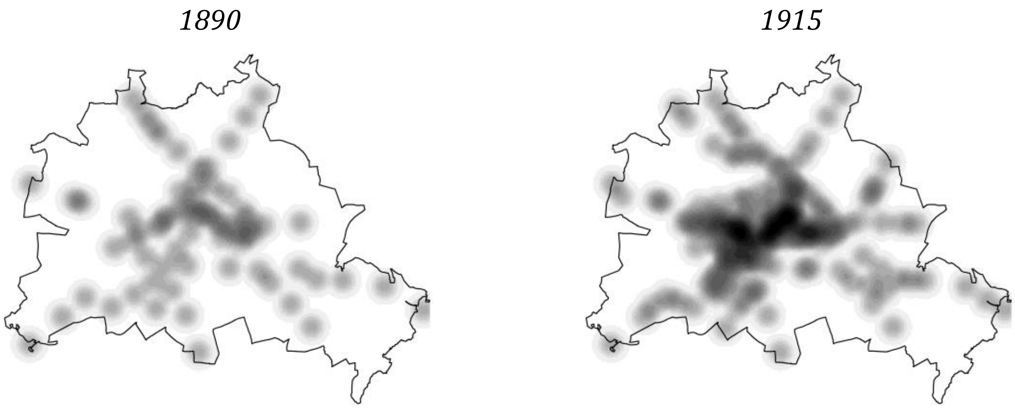
[8] The kernel is defined as $\frac{3}{\pi h^2}(1-t^2)^2$, for $t=\frac{d}{h}\leq 1$, where d is the distance from the analysis location x to location y, h denotes the bandwidth (radius from analysis point to edge of kernel area); 0 for $t>1$.

Fig. 2. Population density



Notes: Data aggregated to neighbourhoods (Ortsteile)

Fig. 3. Rail station density



Notes: Rail station density is first computed in continuous space using a kernel radius of 2km and then aggregated to a 150x150 meter grid. High densities are dark shaded using a consistent scale for both years.

We combine the raw data shown in Figures 1–3 to three panel data sets for the PVAR analyses. First, a bivariate data set containing grid-cell level (150x150m) land values and station densities for six periods (1890–1914) in approximately five-year intervals. Second, a neighbourhood level (48 spatial unit) data set containing land values, rail densities and population for six periods (1890–1914). Third, we also construct a neighbourhood level (93 spatial unit) data set containing rail densities and population for 14 periods (1870–1935). To keep the presentation compact we relegate the analysis of this data set to the appendix. Table 1 provides descriptive statistics of the three variable balanced panel data.

In various robustness checks, we split the area into a core (primarily commercial) and a periphery (primarily residential) area in order to control for the diverging roles of central and peripheral areas. In doing so we adhere to the historic definition of the CBD provided by Leyden (1933). A final data set utilised in a complementary land use analysis consists of a set of maps showing real land uses, which were digitised in a procedure similar to the land value extraction (Aust, 1986).

Tab. 1. Descriptive statistics

Year	Number of Stations (S-/U-Bahn)	Land value Mean (S.D.)	Min Max	Population Mean (S.D.)	Min Max	Station density Mean (S.D.)	Min Max
1890	65	113.680	2.510	38336.7	0	0.212	0.006
(1890)	(65/0)	(175.395)	844.671	(51894.2)	201681	(0.159)	0.424
1895	88	150.107	3.624	43970.2	0	0.247	0.008
(1896)	(88/0)	(224.745)	1032.744	(56294.5)	213384	(0.164)	0.490
1900	96	174.318	2.659	52531.7	0	0.255	0.008
(1900)	(96/0)	(251.755)	1112.501	(64722.3)	253149	(0.162)	0.490
1905	118	197.684	3.268	61719.1	0	0.297	0.008
(1904)	(103/15)	(277.079)	1206.241	(73030.2)	277095	(0.187)	0.607
1910	137	236.704	4.138	70726.4	1457	0.411	0.008
(1910)	(109/28)	(327.020)	1343.457	(84175.4)	309551	(0.311)	0.983
1915	155	181.394	3.673	70973.0	1431	0.524	0.008
(1914)	(109/46)	(232.511)	935.405	(85554.2)	313826	(0.427)	1.228

Note: Years in parenthesis refer to availability of land value data. All variables except the number of stations are in logs. Table shows raw values before time-demeaning and forward-mean-differencing.

3 Results

In the presentation of the results we focus on two of the three models we introduced in section 2. In section 3.1 we analyse the interaction between land value, an independent land use measure of transport demand, and transport supply using a bivariate land value station density PVAR model. In section 3.2 we present estimates of a multivariate population density land value station density PVAR model to gain additional insights into how the relationship between transport demand and supply is driven by land use (commercial vs. residential). To save space, the presentation of the third model discussed in section 2, the bivariate population density station density model, is largely contained to the appendix. A brief qualitative summary of the results of this model is presented at the end of section 3.2. Moreover, we refer to both bivariate models (land value and population density vs. station density) in section 3.3 where we present a quantitative interpretation that connects to the previous literature and we discuss the nature of the PVAR results relative to alternative estimation approaches.

We note that all of our time-demeaned and Helmert transformed (forward-mean-differenced) variables (population, land value, rail density) pass the first generation unit root tests. Population and rail density also pass the second generation unit root test allowing for cross-sectional dependency. We cannot apply the second generation unit root test (Pesaran, 2007) to the land value panel due to the limited number of consecutive periods. A more detailed discussion of the unit root test results is in section 4 in the technical appendix.

3.1 Land value: Bivariate demand supply models

Table 2 displays the reduced form results of the bilateral relationship between land value and effective rail network accessibility. The results are in line with a mutual dependence of economic impact and transport supply as expected theoretically. Past realisations of transport supply positively impact on the contemporary land values, which are a capitalisation of productive land use, and vice versa.

Tab. 2. Reduced form: Bivariate land value model (150m grid level)

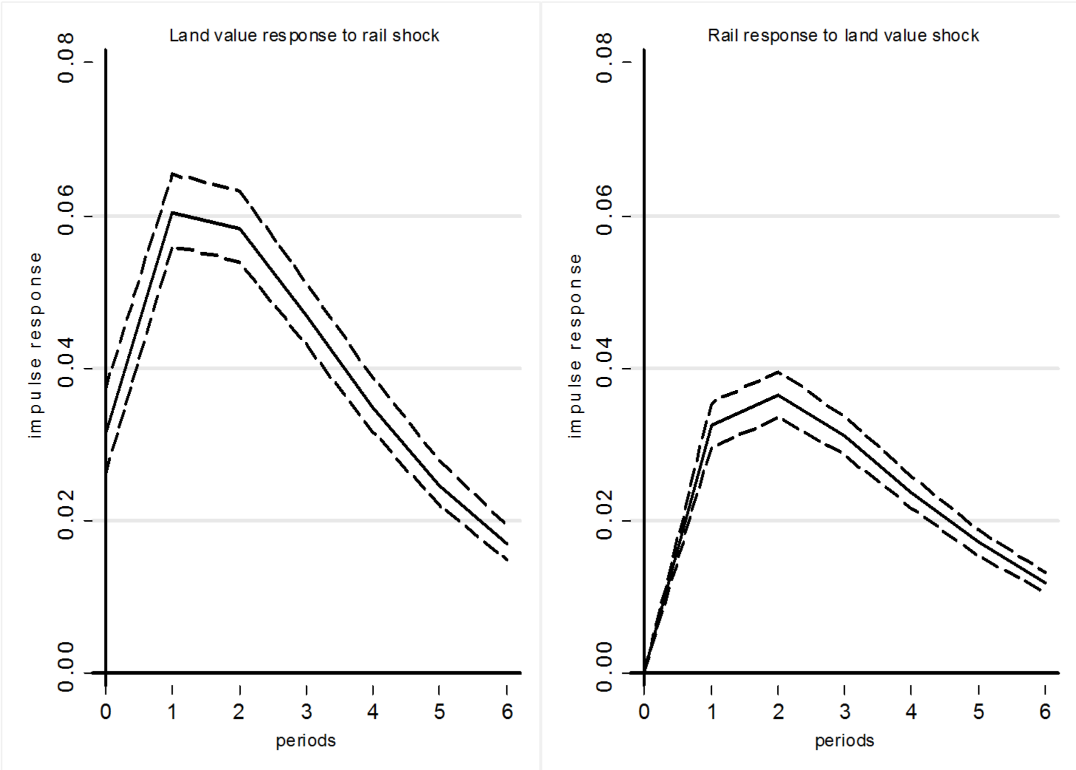
	(1) log land value (t)		(2) log station density (t)	
	Coeff.	S.E.	Coeff.	S.E.
log land value (t-1)	0.590***	(0.008)	0.053***	(0.003)
log station density (t-1)	0.099***	(0.006)	0.539***	(0.014)
Obs.	34,244			

Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Figure 4 illustrates the computed IRF based on the reduced form results presented in Table 2. The IRF summarise the temporal response pattern for a given variable to shocks in the other. The IRF are in line with the positive mutual dependency suggested by the reduced form results. Positive shocks to transport supply lead to positive adjustments in land values in a number of consecutive periods. The IRF converges towards zero without becoming negative, which implies that transport improvements lead to permanent level shifts in the intensity of land use. Over the six time periods the (log) land value adjustment to a one standard deviation shock in log station density accumulates to about 0.27 standard deviations (cumulative IRF are presented in Figure A3 in the appendix). This pattern is in line with a relatively large literature that has provided evidence of a causal impact of transport infrastructure on real estate prices. The station density IRF, however, also indicates that positive economic shocks lead to increases in the supply of transport services.

Note that contemporary shocks are ruled out mechanically. The plateau from the first to the second period indicates that, in practice, it takes some time for transport infrastructure to fully adjust to demand shocks. This is comprehensive in light of the intensity of the heavy rail planning and construction process.

Fig. 4. Impulse responses: Bivariate land value model (150m grid level)



Notes: IRF illustrate the effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation. Dashed lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

Compared to the effect of transport supply on land value, the reverse impact is somewhat smaller. The IRF of station density (Figure 4, right) shows a weaker amplitude than the land value IRF (left). A one standard deviation shock in (log) land value increases (log) station density by about 0.037 standard deviations in the second consecutive period, while similar station density shocks lead to land value responses of about 0.06 (again, in units of standard deviation). Also, the cumulative impulse response of supply to the demand shock is just about half the size of the demand response to the supply shock (0.14 standard deviations over the six periods). This is not surprising as the supply of transport facilities is costly and is usually a matter of political dispute. We would not expect every economically successful neighbourhood to instantly receive a train station. Finally, a variance decomposition analysis[9] (Table 3) indicates that the percentage of variation in station density explained by land value (0.5%) is just about half the magnitude of the share of variation in land value explained by station density (0.9%).

Tab. 3. Bivariate land value model: Variance decomposition

Percent of variation in		log land value	log station density
Explained by	log land value	0.991	0.005
	log station density	0.009	0.995

Notes: Reduced form results (IRF) in Table 2 (Figure 4).

Overall these results support the theoretically expected mutual dependency of economic output and transport infrastructure. While not unexpected, these results also have important implications for the growing literature that aims at estimating the causal impact of transport supply on economic output. The endogeneity of transport supply to transport demand empirically demonstrates the importance of identifying the economic impact of transport investments from exogenous variation in transport supply.

To evaluate the robustness of the benchmark results presented and discussed in the previous section we have altered these models along a variety of dimensions. This section summarises the results of these complementary analyses. Detailed discussions are in section 5 of the appendix.

Modifiable areal unit problem (MAUP)

A common concern in many spatial analyses is that the level of spatial aggregation of the underlying data may affect the results. This concern is typically directed towards arbitrarily defined boundaries of official statistical units such as for our neighbourhoods used in the multivariate PVAR model. We are not able to alter the unit of analysis due to a lack of data in this model. However, we are able to evaluate the extent to which aggregation to larger grid cells affects the results in the bivariate land value station density models. Aggregation results in very similarly shaped IRF (Figure A4).

[9] More detailed variance decompositions for the geographical subsamples can be found in section 6 of the appendix.

Transport accessibility measurement

Our baseline transport access measure assumes that the impact of an additional station in a neighbourhood does not depend on the number of stations already there, which imposes a strong form of complementarity of the services offered. We have replicated our benchmark models using the distance to the nearest station, which is a popular measure in the literature and imposes that stations are perfect substitutes. The results remain qualitatively and quantitatively similar (Figure A5).

Announcement network

Under the weak-form efficient market hypothesis we expect that asset prices should incorporate all available information that is publicly available, which implies the capitalisation of new transport infrastructures at the time of announcement (not completion). To accommodate such effects we rerun our benchmark models using historical station density measures based on the dates of announcements. While the shape of the IRF, not surprisingly, changes somewhat, all of the qualitative conclusions from the benchmark models still apply (Figure A6).

Complementary IV analysis

An important and policy relevant conclusion from the PVAR analysis presented so far is that transport supply shocks have a positive and permanent impact on the productivity of land use reflected in land value An interesting question is how these results generated by the novel application of the PVAR method compare to results produced by standard techniques of causal inference. To answer the question we have estimated the impact of changes in station density on land value using the more established panel IV method. Our models are estimated in differences to remove unobserved spatial heterogeneity, allow for heterogeneous long-run trends at the plot level, and control for unobserved macroeconomic shocks at the neighbourhood level. Following Gibbons et al. (2012) we argue that a quasi-experimental variation in transport supply on a fine geographical level is as good as random because the routing then becomes exogenous.

To further strengthen our identification we use an IV to restrict the variation in station density used for identification to a fraction that we argue to be exogenous. Therefore we compute a station density measure based on a counterfactual heavy rail transport network used by (Ahlfeldt & Wendland, 2011). The network consists of straight lines that connect the CBD to the most important nearby towns as well as an emerging secondary centre (the Kurfürstendamm). We distribute counterfactual stations every 1,089 metres along the IV tracks, where 1,089 metres is the average distance between rail-bound stations in 1915. We run locally weighted regressions (LWR) (Cleveland & Devlin, 1988; McMillen, 1996) of actual densities on the counterfactual density for each period and recover the predicted values, which form our time-varying IV. To rationalise this strategy, we argue that being closer to the potential transport corridors increases the chance of being connected to the network over the study period. At the same time, being closer to the hypothetical network the conditional *neighbourhood x period* effects is as good as random and there is little reason why any (conditional) temporal trend correlated with this measure should exist for reasons other than the improvements in transport services supply we are interested in. Briefly summarised, we find positive and significant transport supply effects on land value, which, as we will discuss in section 3.3, are quantitatively close to the PVAR estimates. A detailed discussion of our complementary land value IV estimation is in section 7 of the appendix.

3.2 Land use: Multivariate demand supply models

Table 4 presents the reduced form results of the three variable PVAR model including population. To allow for a structural interpretation, the IRF are displayed in Figure 5. Compared to the bivariate model discussed above, the three variable model allows for additional insights into the effects transport supply shocks exert on land use pattern and vice versa.

The population response to rail shocks (bottom right) is significantly negative over a number of consecutive periods. Under the assumption of market clearing, a negative population response must correspond to a positive response in the competing commercial land use. The implication is that improvements in transport access tend to attract firms at the expense of displacing residents. Likewise, the negative response of transport supply (bottom middle) to population shocks, in a world without perfectly elastic floor space, can be interpreted as a positive supply reaction to positive employment shocks that displace population. Thus, above and beyond a general supply response to increases in locational productivity reflected in land value shocks (upper middle), changes from residential to commercial land use lead to improvements in transport access. The implication is that the transport planner targets commercial transport demand in particular, which seems sensible in light of the transport *gravity*. A higher density of economic activity potentially attracts customers and employees from across the urban area and generates potentially larger local transport demand than pure residential use.

The population response to land value shocks, all else equal, is positive and just about significant (upper-right). Shocks that make a location fundamentally more attractive (higher land value) attract residents to that location. Similarly, positive population shocks capitalise into land values (bottom-left). At the same time, the land value response to population shocks is flat and is not statistically distinguishable from zero (upper-left). Taking together the insignificant *conditional* land value response to rail shocks (Figure 5, upper-left) and the positive *unconditional* land value response to transport shocks (Figure 4, left) we conclude that the positive effect of a transport improvement on the value of land operates primarily through increased residential demand for more accessible space.

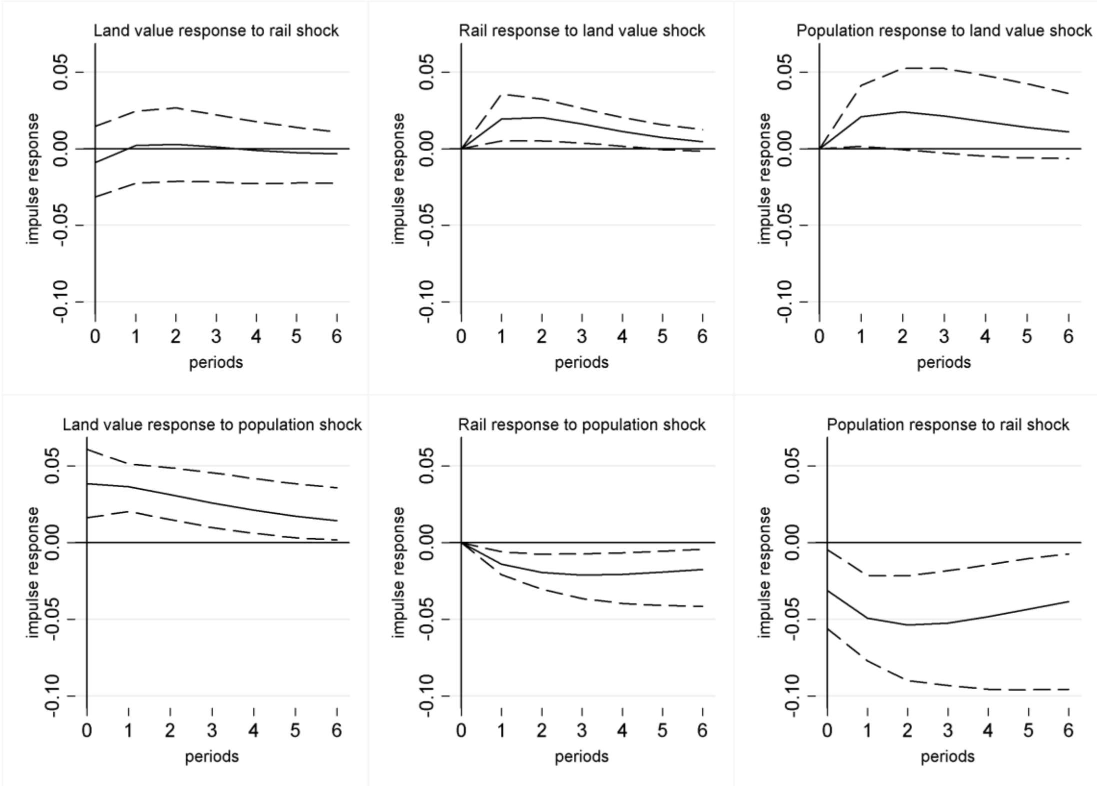
As in the bivariate models the variance decomposition (see Table 5) indicates that the presumably most exogenous variable – transport supply – explains a significantly larger fraction of the most endogenous variable – land value – than the other way round.

Tab. 4. Reduced form: Multivariate land use model (neighbourhood level)

	(1) log land value (t)		(2) log station density (t)		(3) log population (t)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
log land value (t-1)	0.500***	(0.075)	0.104***	(0.048)	0.112*	(0.067)
log station density (t-1)	0.057	(0.066)	0.639***	(0.121)	-0.146***	(0.069)
Log population (t-1)	0.084***	(0.037)	-0.088***	(0.027)	0.784***	(0.069)
Obs.			188			

Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. 4 out of 48 Ortsteile were not incorporated as they had zero population until the end of our observation period and are treated as missing values.

Fig. 5. Impulse responses: Multivariate land use model (neighbourhood level)



Notes: IRF illustrate the effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation. Dashed lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

Tab. 5. Multivariate land use model: Variance decomposition

Percent of variation in		log land value	log population	log station density
Explained by	log land value	0.938	0.060	0.002
	log population	0.006	0.947	0.047
	log station density	0.010	0.005	0.985

Notes: Reduced form results (IRF) in Table 4 (Figure 5).

Both the bivariate and the multivariate PVAR model support the mutually dependent relationship between our economic outcome measure land value and our transport supply measure station density. The multivariate model further reveals that transport improvements can be the cause and effect of land use changes. Improvements in transport tend to lead to an outbidding of residents by firms. Likewise, a switch from residential to commercial use tends to improve transport supply, indicating that the transport planning process targets firms more than residents. Jointly, the bivariate and multivariate model results suggest that on average the transport effect on land value is primarily moderated by an increased residential demand for accessible locations.

As with the bivariate demand and supply model we have conducted a number of complementary robustness checks. While we only discuss models that address land use-related issues here we have also conducted a number of additional robustness checks that address other concerns (e.g. measurement of transport access, announcement effects). All robustness checks are presented in section 6 in the appendix.

Land use: Bivariate demand supply model

We finally estimate an alternative bivariate model which uses population density as a measure of economic activity instead of land values (Levinson, 2008). In this model the analysis then needs to be disaggregated by land use. The reason is that depending on the relative attractiveness of an area to firms and residents, the population will either be attracted to an area with improved transport supply, or displaced by firms that relocate to the area. We have therefore split our study area into a core (primarily commercial) and a periphery (primarily residential) and replicated our analyses using these spatial subsamples.

The models presented in the previous sections are somewhat constrained in their spatial and temporal coverage by the availability of land value data. Our population data ranges from 1870 to 1935 and covers an area that roughly corresponds to today's Berlin. We run bivariate demand supply models for the subsample used in benchmark models (1890–1915), the whole data set (1870–1935) as well as the core and the periphery area separately. Expanding the spatial and temporal scope of the analysis only marginally altered the results. Similar to Levinson (2008) we find that the relationship between population and rail density are qualitatively distinct in the core and periphery sample. While new infrastructure tends to displace residents in the core area, it attracts residents in the peripheral area. The bivariate land use models are presented in detail in section 6.4 of the appendix.

Complementary IV analysis

To evaluate the main finding of this section—that transport supply shocks lead to conversion of residential into commercial use—we use a complementary IV approach similar to the one discussed at the end of section 3.1. While we cannot explicitly compare the magnitudes of the implied land use changes across estimation techniques, it is notable that our analysis of real land use changes clearly indicates that increases in station density lead to a conversion of residential to commercial use, especially in central areas, which is in line with the conclusions we have derived from the PVAR models. The analysis is presented in detail in section 7 of the appendix.

3.3 Quantitative interpretation

In this section we present some back-of-the-envelope calculations that help to express the results from the bivariate PVAR analysis in intuitive metrics that can be compared to previous research and alternative methods.

We begin with the relationship between transport demand as measured by population density and transport supply as measured by station density, which was previously analysed by Levinson (2008), henceforth L. For a quantitative comparison we focus on the effect of station density on population in the peripheral sample of neighbourhoods in London, for which L reports significant effects (unlike for the core). L reports a semi-elasticity of population with respect to station density of 0.33 (L, Table 3). At a mean station density of 0.1 (L, Figure 2), the implied elasticity of population with respect to station density is 3.3% (=0.33/0.1). This is a short-run effect over one period that corresponds to 10 years. In our bivariate population station density PVAR model estimated for the periphery sample (lower-left panel in Fig. A13 in the appendix) we find that in the short run (one period corresponding to five years) a one SD increase in log station density increases log population by 0.04 SD. Given an SD in log station density of 1.23 and an SD in log population of 0.84 in our data the estimated effect corresponds to an elasticity of population with respect to station density of 2.7% (=0.04 × 0.84 / 1.23). Our estimates are thus not only qualitatively but also quantitatively similar to L. The somewhat lower implied elasticity is probably at least partially attributable to the shorter time-frame of our short-run effect (5 vs. 10 years).

Next, we turn our attention to the relationship between transport supply and land price. A number of studies that have used cross-sectional hedonic methods or quasi-experimental designs such as DiD tend to report the percentage effects of one unit reductions in distance to a station on land prices or house prices. As an example, Gibbons & Machin (2005), henceforth GM, in one of the most careful analyses of the house price effects of mass transit report a 1.5–5.5% effect caused by a 1km reduction in distance to the nearest metro rail station following the 1999 extension of the Jubilee Line and the DLR in London, UK. Ahlfeldt & Wendland (2009), henceforth AW, report a 2–2.5% effect of a 100m reduction in station distance on land values for Berlin (1890-1936). To allow for a rough comparison between GM's reported effects on house prices and AW's as well as our estimated land price effects it is useful to assume a Cobb-Douglas housing production function and a competitive construction sector. Following Combes et al. (2013) it is then possible to translate a house price effect into a land price effect by dividing the former by a land share parameter of 0.25.[10] GM's results then correspond to a land price effect of 0.6–2.2% per 100m reduction in station distance, which is reasonably close to AW's findings.

In an attempt to make both GM's and AW's results comparable to our findings we now express all estimates as elasticity at the mean. The mean change in station distance in the group of treated properties reported by GM is -1,558m (GM, Table 1, column 7). For such a change in station distance the implied elasticity of land price with respect to station distance is –(9.3%-34.3%) (= -[0.6%-2.2%] × -1,558 / -100). The mean change in distance to station in AW's data is -550m (own analysis of their data set). The implied elasticity of land price with respect to station distance is thus –(11-13.8%) (= -[2-2.5%] × -550 / -100). Because GM and AW identify their effects from a 2km station catchment area and because the kernel in our station density measure uses a radius of 2km it is reasonable to assume that the magnitudes of changes in station distance and station density following transport innovations are approximately proportionate. It is therefore reassuring that the results from our IV analysis (≈13%) are extremely close to AW, who analyse a small subsample of our data, and within the range of GM.

In our bivariate land value station density model, we estimate a short-run (i.e. five years) land value increase of 0.06 SD in response to a transport shock (Fig. 4, left panel). Given an SD in log station density of 0.77 and an SD in log land values of 0.58 in our data the estimated effect translates into an elasticity of 8% (= 0.06 x 0.77 / 0.58), which is only slightly below the implied elasticity effects in GM and AW. The cumulated long-run effect (six periods, 30 years) of 0.27 SD (Fig. A3) similarly corresponds to an elasticity of land price with respect to a transport supply of 36%. Applying the same transformation suggested by

[10] Assume that housing services H are produced using the inputs capital K and land L as follows : $H = K^{\alpha} L^{1-\alpha}$. Housing space is rented out at bid-rent ψ while land is aquired at land rent Ω. From the F.O.C. $K/L = \alpha/(1-\alpha)$ (the price of capital is the numeraire) and the non-profit condition $\psi H = K + \Omega L$ it follows imediately that $\log(\psi) = 1/(1-\alpha)\log(\Omega) + c$, where c is some unimportant constant that cancels out in first-differences, i.e. $\Delta\log(\psi) = 1/(1-\alpha)\Delta\log(\Omega)$.

Combes et al. (2013) as above, the implied elasticity of house price with respect to station density is 2% (short run) to 8.5% (long run). A tentative comparison of our PVAR estimates to GM, AW and our own panel IV/DiD estimates suggest that the DiD method tends to deliver estimates that are closer to our short-run than long-run PVAR estimates.

4 Conclusion

With this contribution, we provide a novel analysis of the simultaneity of the supply of and demand for transportation, which resembles the well-known chicken-and-egg problem. We borrow a method from the macro-econometric literature designed to empirically approximate systems of mutually dependent variables: structural PVAR modelling.

We argue that assuming market clearing and a competitive construction sector with a concave production function, the relationship between transportation demand and supply can be approximated based on a measure of transport access and a measure of land value, which reflects the economic productivity of the land. Adding a measure capturing the intensity of residential (or commercial) use to a PVAR system, it is further possible to conclude on how land use determines transport supply and vice versa.

Our results confirm the long suspected mutually dependent relationship between local economic performances on the one hand and transportation supply on the other. An exogenous increase in the effective supply of transport services by one SD, all else equal, leads to a short-run adjustment in land values of about 0.06 SD and a cumulated effect of 0.27 SD after six periods. The implied short-run (long-run) elasticity of property price with respect to transport supply is 2% (8.5%). The reverse effect of an exogenous increase in economic activity that increases the value of land by one SD leads to a positive, but significantly lower 0.14 SD effect on transport supply. Moreover, transport supply shocks seem to trigger a displacement of residential for commercial land use. On the supply side, shifts from residential to commercial land use attract new transport supply. These results are robust to different measures of transport access, different forms of assumed market efficiency that determine the timing of capitalisation into land values, and are consistently found within the commercial core and the residential periphery of the city.

Importantly, our results from the PVAR analysis are also qualitatively (land use) or even quantitatively (land value) consistent with the findings from a complementary analysis that uses established toolkits of causal unidirectional inference. The DiD estimates we reviewed tend to be closer to our short-run than long-run PVAR estimates. Notwithstanding these encouraging results we do not propose using the PVAR method as a replacement for conventional identification strategies when it comes to the impact analysis of transport projects. We do, however, argue that the PVAR method is a useful complementary tool that produces a generic picture of the dynamics of transportation demand and supply, which covers the multi-directionality of impact, the temporal structure of the dependencies and even the interactions with land use.

Literature

Ahlfeldt, G. M. (2013). If we build it, will they pay? Predicting property price effects of transport innovations. *Environment and Planning A, 45*(8), 1977 – 1994.

Ahlfeldt, G. M., & Wendland, N. (2009). Looming Stations: Valuing Transport Innovations in Historical Context *Economics Letters, 105*(1), 97-99. doi: http://dx.doi.org/10.1016/j.econlet.2009.06.010

Ahlfeldt, G. M., & Wendland, N. (2011). Fifty Years of Urban Accessibility: The Impact of Urban Railway Network on the Land Gradient in Berlin 1890-1936. *Regional Science and Urban Economics, 41*(2), 77-88.

Ahlfeldt, G. M., & Wendland, N. (2013). How polycentric is a monocentric city? Centers, spillovers and hysteresis. *Journal of Economic Geography, 13*(1), 53-83. doi: 10.1093/jeg/lbs013

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

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies, 58*, 277-297.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error component models. *Journal of Econometrics, 68*, 29-51.

Assenmacher-Wesche, K., & Gerlach, S. (2008). Monetary policy, asset prices and macroeconomic conditions : a panel-VAR study: National Bank of Belgium.

Aust, B. (Cartographer). (1986). Die städtebauliche Entwicklung Berlins.

Baum-Snow, N. (2007). Did Highways Cause Suburbanization? *The Quarterly Journal of Economics, 122*(2), 775-805. doi: 10.1162/qjec.122.2.775

Baum-Snow, N., Loren Brandt, J., Henderson, J. V., Turner, M. A., & Zhang, Q. (2012). Roads, Railroads and Decentralization of Chinese Cities. *Working Paper*.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87*, 115-143.

Bogart, D. (2009). Inter-Modal Network Externalities and Transport Development: Evidence from Roads, Canals, and Ports During the English Industrial Revolution. *Networks and Spatial Economics, 9*(3), 309-338. doi: 10.1007/s11067-008-9060-0

Brady, R. R. (2011). Measuring the diffusion of housing prices across space and over time. *Journal of Applied Econometrics, 26*(2), 213-231. doi: 10.1002/jae.1118

Brueckner, J. K. (1987). The structure of urban equilibria: A unified treatment of the Muth-Mills model. In E. S. Mills (Ed.), *Handbook of Regional and Urban Economics* (Vol. 11, pp. 821-845). Amsterdam: North-Holland.

Calderón, C., Moral-Benito, E., & Servén, L. (2014). Is infrastructure capital productive? A dynamic heterogeneous approach. *Journal of Applied Econometrics*, n/a-n/a. doi: 10.1002/jae.2373

Carlino, G., & DeFina, R. (1998). The differential regional effects of monetary policy. *Review of Economics & Statistics, 80*(4), 572-587.

Cervero, R., & Hansen, M. (2002). Induced Travel Demand and Induced Road Investment – A Simultaneous Equation Analysis. *Journal of Transport Economics and Policy, 36*(3), 469-490.

Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance, 20*, 249-272.

Cleveland, W. S., & Devlin, S. J. (1988). Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association, 83*(403), 596-610.

Combes, P.-P., Duranton, G., & Gobillon, L. (2013). The Costs of Agglomeration: Land Prices in French Cities. *Paris School of Economics Working Paper No 2013 - 28*.

Duranton, G., & Turner, M. A. (2011). The Fundamental Law of Road Congestion: Evidence from US Cities. *American Economic Review, 101*(6), 2616-2652. doi: 10.1257/aer.101.6.2616

Duranton, G., & Turner, M. A. (2012). Urban Growth and Transportation. *The Review of Economic Studies, 79*(4), 1407-1440. doi: 10.1093/restud/rds010

Enders, W. (1995). *Applied econometric time series*: Wiley.

Epple, D., Gordon, B., & Sieg, H. (2010). A New Approach to Estimating the Production Function for Housing. *American Economic Review, 100*(3), 905-924. doi: doi: 10.1257/aer.100.3.905

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance, 25*(2), 383-417.

Fujita, M., & Ogawa, H. (1982). Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional Science and Urban Economics, 12*(2), 161-196.

Gibbons, S., Lyytikäinen, T., Overman, H., & Sanchis-Guarner, R. (2012). New Road Infrastructure: the Effects on Firms. *SERC Discussion Paper 117*.

Gibbons, S., & Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics, 57*(1), 148-169.

Gillanders, R. (2011). The Effects of Foreign Aid in Sub-Saharan Africa: School Of Economics, University College Dublin.

Graham, D. J., Melo, P. S., Jiwattanakulpaisarn, P., & Noland, R. (2010). Testing For Causality Between Productivity And Agglomeration Economies. *Journal of Regional Science, 50*(5), 935-951.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica, 37*(3), 424-438.

Gravier-Rymaszewska, J. (2012). How Aid Supply Responds to Economic Crises: A Panel VAR Approach: World Institute for Development Economic Research (UNU-WIDER).

Hamilton, J. D. (1994). *Time Series Analysis*: Princeton University Press.

Holl, A., & Viladecans-Marsal, E. (2011). Infrastructure and Cities: The Impact of New Highways on Urban Growth. *Universitat de Barcelona-Institut d'Economia de Barcelona Working Paper*.

Holtz-Eakin, D., Newey, W., & Rosen, H. S. (1988). Estimating Vector Autoregressions with Panel Data. *Econometrica, 56*(6), 1371-1395.

Hornung, E. (2012). Railroads and Micro-regional Growth in Prussia: Ifo Institute for Economic Research at the University of Munich.

Hsu, W.-T., & Zhang, H. (2011). The Fundamental Law of Highway Congestion: Evidence from Japanese Expressways. *Chinese University of Hong Kong Working Paper*.

Inoue, A., Kilian, L., & Kiraz, F. B. (2009). Do Actions Speak Louder Than Words? Household Expectations of Inflation Based on Micro Consumption Data. *Journal of Money, Credit and Banking, 41*(7), 1331-1363.

Kalweit, F. (1928). *Die Baustellenwerte in Berlin*. Berlin: Ermo Verlagsgesellschaft.

Kalweit, F. (1936). *Die Baustellenwerte in Berlin*. Berlin: Carl Heymanns Verlag.

Kilian, L. (2011). Structural Vector Autoregressions: C.E.P.R. Discussion Papers.

Konstantinos, D., & Konstantinou, P. T. (2013). Terrorism Shocks and Public Spending: Panel VAR Evidence from Europe. *Defence and Peace Economics*.

Lee, C.-I. (2007). Does provision of public rental housing crowd out private housing investment? A panel VAR approach. *Journal of Housing Economics, 16*(1), 1-20.

Levinson, D. (2008). Density and dispersion: the co-development of land use and rail in London. *Journal of Economic Geography, 8*(1), 55-77.

Levinson, D., & Karamalaputi, R. (2003). Induced Supply: A Model of Highway Network Expansion at the Microscopic Level. *Journal of Transport Economics and Policy, 37*(3), 297-318.

Leyden, F. (1933). *Gross-Berlin. Geographie der Weltstadt*. Breslau: Ferdinand Hirt.

Love, I., & Zicchino, L. (2006). Financial development and dynamic investment behavior: Evidence from panel VAR. *The Quarterly Review of Economics and Finance, 46*(2), 190-210.

Lucas, R. E., Jr., & Rossi-Hansberg, E. (2002). On the Internal Structure of Cities. *Econometrica, 70*(4), 1445-1476.

M'Amanja, D., & Morrissey, O. Foreign Aid, Investment and Economic Growth in Kenya: a Time Series Approach: University of Nottingham, CREDIT.

Mauruszat, A. (2010). Bahnstrecken im Land Brandenburg - interaktive Karte Retrieved February 25th, 2011, from http://www.bahnstrecken.de

McDonald, J. F., & Osuji, C. I. (1995). The effect of anticipated transportation improvement on residential land values. *Regional Science & Urban Economics, 25*(3), 261.

McMillen, D. P. (1996). One Hundred Fifty Years of Land Values in Chicago: A Nonparametric Approach. *Journal of Urban Economics, 40*(1), 100-124.

McMillen, D. P., & McDonald, J. F. (2004). Reaction of House Prices to a New Rapid Transit Line: Chicago's Midway Line, 1983-1999. *Real Estate Economics, 32*(3), 463-486.

Melo, P. S., Graham, D. J., & Canavan, S. (2012). Effects of Road Investments on Economic Output and Induced Travel Demand Evidence for Urbanized Areas in the United States. [Article]. *Transportation Research Record*(2297), 163-171. doi: 10.3141/2297-20

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

Miller, N., & Peng, L. (2006). Exploring Metropolitan Housing Price Volatility. *The Journal of Real Estate Finance and Economics, 33*(1), 5-18.

Mills, E. S. (1972). *Studies in the Structure of the Urban Economy*. Baltimore: Johns Hopkins Press.

Muth, R. F. (1969). *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. Chicago: University of Chicago Press.

Osei, R., Morrissey, O., & Lloyd, T. (2005). The fiscal effects of aid in Ghana. *Journal of International Development, 17*(8), 1037-1053.

Pereira, A. M., & Andraz, J. M. (2012a). On the effects of highway investment on the regional concentration of economic activity in the USA. *Portuguese Economic Journal, 11*(3), 165-170. doi: 10.1007/s10258-012-0080-9

Pereira, A. M., & Andraz, J. M. (2012b). On the regional incidence of highway investments in the USA. *The Annals of Regional Science, 48*(3), 819-838. doi: 10.1007/s00168-010-0409-1

Pereira, A. M., & de Frutos, R. F. (1999). Public Capital Accumulation and Private Sector Performance. *Journal of Urban Economics, 46*(2), 300-322. doi: http://dx.doi.org/10.1006/juec.1998.2124

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics, 22*(2), 265-312.

Runge, E. (Cartographer). (1950). Bodenwertkarte Gross-Berlin (1938).

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics, 125*(3), 1253-1296. doi: 10.1162/qjec.2010.125.3.1253

Schomacker, M. (2009). Berliner-Untergrundbahn.de Retrieved February 25th, 2011, from http://www.berliner-untergrundbahn.de/

Silverman, B. W. (1986). *Density Estimation*. London: Chapman and Hall.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica, 48*(1), 1-48.

Statistisches Amt der Stadt Berlin. (1920). *Statistisches Jahrbuch der Stadt Berlin*. Berlin.

Straschewski, M. (2011). Geschichte und Geschichten rund um die Berliner S-Bahn Retrieved February 25th, 2011, from http://www.stadtschnellbahn-berlin.de

Xie, F., & Levinson, D. (2010). How streetcars shaped suburbanization: a Granger causality analysis of land use and transit in the Twin Cities. *Journal of Economic Geography, 10*(3), 453-470.

Gabriel M. Ahlfeldt[♣], Kristoffer Moeller[⊗] & Nicolai Wendland[♦]

Technical appendix to Chicken or egg? The PVAR econometrics of transportation

1 Introduction

This appendix complements the main paper and is not designed to stand alone or replace the main paper. Sections 2 and 3 provide additional background on the development of the rail system in Berlin and the data used for the PVAR analyses. Section 4 provides the results of unit root tests which were carried out prior to the actual regressions. Section 5 complements the bivariate land value model results presented in section 3.1 of the main paper by providing complementary results and robustness tests. Section 6 similarly presents complementary results and a range of robustness checks for the multivariate land use model discussed in section 3.2 in the main paper. The section also discusses a bivariate model that shares similarities with Levinson (2008). Finally, section 7 presents a complementary and independent panel IV analysis of the impact of transport infrastructure on land value and land use which we use to benchmark the interpretations derived from the PVAR models.

2 Background

The public rail network in Berlin is made up of two different modes, namely a suburban rail system and the underground. We give a short historic overview of its development in this chapter before giving a brief description of our data.

♣ *Corresponding author:* London School of Economics and Political Sciences & Spatial Economics Research Centre (SERC). g.ahlfeldt@lse.ac.uk, www.ahlfeldt.com

⊗ Darmstadt University of Technology and Center for Metropolitan Studies TU Berlin mail@kristoffer-moeller.de, www.kristoffer-moeller.de.

♦ Darmstadt University of Technology. wendland@vwl.tu-darmstadt.de

2.1 S-Bahn network

The suburban rail ("S-Bahn") as it is known today is a result of combining various suburban lines ("Vorortsbahn"), the original city line ("Stadtbahn") and the circular line ("Ringbahn") in 1930. As a result, there are various reasons as to why, where and how the S-Bahn was developed over the years which originate from the three different strands (Gottwaldt, 1994; Kiebert, 2004, 2008; Klünner, 1985):

The suburban lines connected Berlin with its surrounding cities and suburbs. The early lines in particular originate from long-distance connections to other important cities like Potsdam ("Stammbahn" 1838), Hamburg ("Hamburger Bahn" 1846) or Dresden ("Dresdner Bahn" 1887), initially sharing their tracks with the new upcoming suburban lines. In 1891 a new tariff system for local mass transit was introduced – pushing the passenger numbers up by about 30% – and the suburban lines increasingly started to run on their own tracks. The majority of these lines were developed by public companies and planned by the government. For instance, the "Ostbahn", which was supposed to go through the Prussian regions of Pommern and East Prussia, was built in order to develop the periphery along the tracks. The "Görlizer Bahn" (1866/67) or the "Wetzlarer Bahn" (also "Canon Train") linking Berlin with Metz at the French border was planned by the military to facilitate the rapid movement of troops. Later on, new lines were built specifically for local mass transit in order to improve the access of the periphery like the North-South connection (1934–39). However, private developers like J. A. W. Carsten, who financed the station "Lichterfelde" (1868) in order to sell his newly established country estates in that area, intervened in the expansion of the S-Bahn network as well. The electronic company Siemens further financially supported the exploitation of the section between Fürstenbrunn and Siemensstadt (1905) in order to improve the commute for its workers. Moreover, the Brothers Spindler, who ran a laundry and dying factory in Köpenick at the Eastern border of the city, lobbied for a transport line between Schöneweide and Spindlersfelde (1891). Hence, the suburban lines were driven by both the public and the private sector.

The city line went from Stralau-Rummelsburg to Westkreuz, Halensee, and was built in 1882. This East-West connection, running through the historical city centre, was planned to decongest the traffic between Berlin's terminal stations. The tracks were mainly built on land owned by the government and the project was carried out publicly.

The first sections of the circular line Moabit-Gesundbrunnen-Potsdamer-Ringbahnhof and Moabit-Charlottenburg-(Westend)-Grunewald-Tempelhof were opened in 1881 and 1882 respectively. The circular line was financed by the state of Prussia but was run by the Niederschlesisch-Märkische Eisenbahn, a public company owned by Prussia. The idea behind the circular line was to connect radian lines going out of the centre with each other and the important terminal stations. Various parts of the new line were built into undeveloped land and thus outside the city border. Or as Elkins & Hofmeister (1988, p. 114) state: "The actual position of the ring line was a compromise between the desire to maximise utilisation by being as close as possible to the core of the city and the desire to minimise land-acquisition costs by avoiding areas of existing urban development."

Like the circular line, many other lines of the light rail system were built into undeveloped areas, connecting Berlin with other villages. Only the East-West and North-South connections went through the city centre. New villages were founded close to the new lines like, for instance, "Glienicke an der Nordbahn". Companies like AEG or Borsigwerke in Tegel even built new factories in close proximity to the new stations ("Kremmener Bahn"). Even though a few S-Bahn lines were developed upon the request of the private sector, most of the lines were developed by the public sector. In the 1880s the majority of the long-distance lines, which were closely related to the rise of the suburban lines, were nationalised. However, most of the nationalised lines were still run independently. They had their own management as well as their own trains/coaches. From 1920 on, all lines were eventually nationalised under the "Reichseisenbahn".

2.2 U-Bahn Network

The underground ("U-Bahn") was developed about a third of a century later than Berlin's light rail system. The first line was opened in 1902 and ran from Stralauer Tor (later Warschauer Brücke) to Potsdamer Platz and then to Zoologischer Garten. The first "underground" section was actually built on elevated tracks since the Berlin government was afraid of damaging its newly installed drainage system. The project was pushed forward by the company "Siemens & Halske", which as early as 1891 had proposed a densely linked network, connecting the historic city centre with its surrounding municipalities. The new line was eventually developed by the "Hochbahngesellschaft", a company jointly founded by Siemens & Halske and Deutsche Bank as the main financer. While the line's Eastern section up to Nollendorfplatz was built on viaducts, the city of Charlottenburg successfully negotiated the tracks to run underground when passing through their territory. Not obstructing the view of the prominent church “Kaiser-Wilhelm-Gedächtniskirche” was one of Charlottenburg’s reasons for the changed routing. In the West (Westend), the line was built into undeveloped land owned by Deutsche Bank. The bank expected to benefit from rising land rents due to the improved access. Followed by the newly established connection Western Charlottenburg turned into an attractive business area. The extension of the first line leading into central Berlin was hampered by the tram operator “Große-Berliner-Straßenbahn” which was afraid of losing its monopolistic role in that area. Eventually, the line ran via Mohrenstraße and Spittelmarkt through the city centre (Gottwaldt, 1994).

The municipalities in the South West showed particular ambition with their plan to develop their unused land. They competed for wealthy citizens by turning it into attractive residential areas. The underground played a crucial role in developing these areas. The city of Schöneberg (“Schöneberger Linie” 1910) even planned and financed its own line between Nollendorfplatz and Hauptstraße (today Innsbrucker Platz) in order to develop its Western territory. As the “Hochbahngesellschaft” did not expect the new line to generate any profits it was completely planned and financed publicly. The land where the lines went through was changed significantly with individually designed stations built at prominent squares. A similar approach was followed by the villages of Wilmersdorf and Dahlem. Newly planned country estates and academic institutes were supposed to benefit from an improved access by constructing the “Wilmersdorf-Dahlemer U-Bahn” (1913). The line was divided into three sections according to their ownerships: While the section between Wittenbergplatz and Nürnberger Platz belonged to the Hochbahngesellschaft, Nürnberger Platz-Breitenbachplatz was owned by the city of Wilmersdorf and Breitenbachpatz-Thielplatz by Domäne Dahlem. The line was extended to the lake “Krumme Lanke” in 1929. This extension was mainly financed by the land speculator and private developer Adolf Sommerfeld in order to connect his newly established residential quarters in Dahlem. Moreover, he wanted to improve the access to the surrounding woods, establishing them as recreational areas (Kurpjuweit & Meyer-Kronthaler, 2009).

In contrast to the S-Bahn network, the initial idea of Berlin’s underground was to serve the local mass transit. The lines were built into more central areas. Moreover the network was developed later than the suburban rail system; the technology was superior, allowing for underground tracks, and both the planners and investors had already gained some initial experience by evaluating the effects of the S-Bahn. Anecdotal evidence suggests that the rise of the U-Bahn was mainly driven by the idea of developing empty land close to the historical core (especially in the South West). Public and private planners alike competed for wealthy citizens and increasing land rents.

Even though the link between transport and land development is not completely clear when analysing the history of Berlin's transport system, the majority of the projects and newly constructed lines appeared to lead the development in an area rather than the other way around.

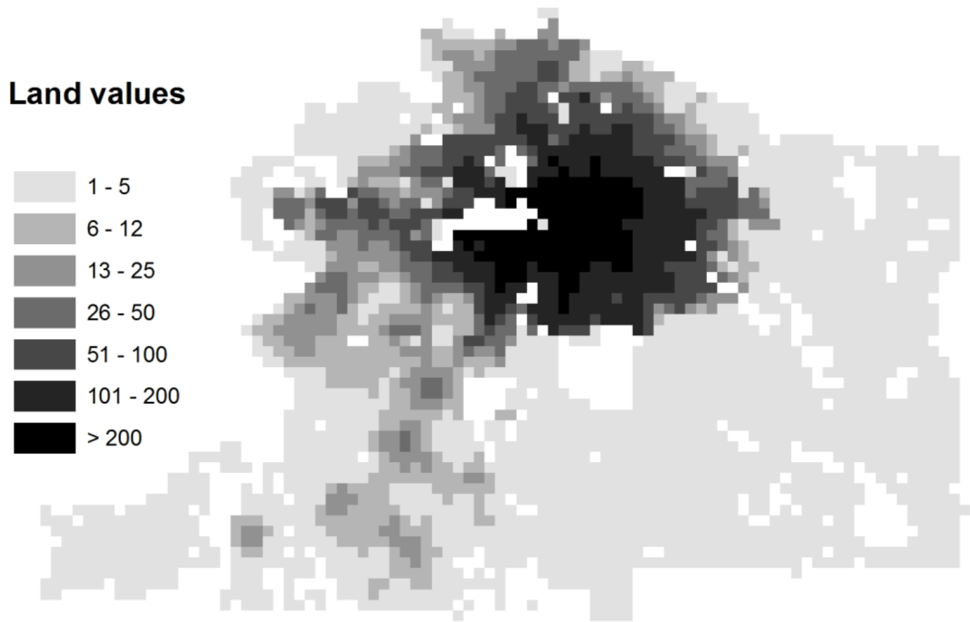
Fig. A1. Rail network in 1880, 1900 and 1915

3 Data

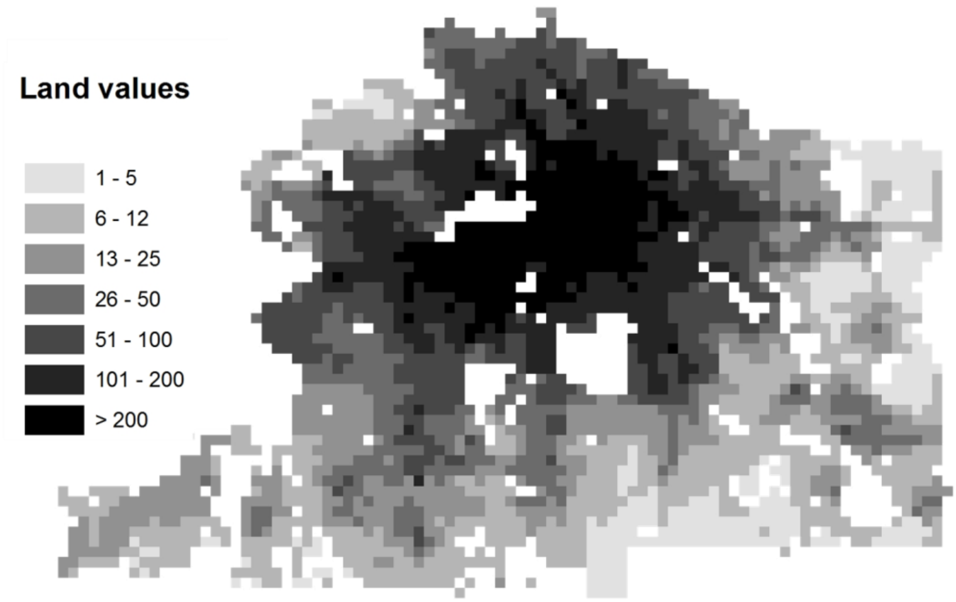
3.1 Separate neighbourhood and grid level models

The different data sources we consider offer information that varies significantly in terms of spatial and temporal coverage. Depending on the data sources we use in our empirical models our analyses are therefore conducted at different levels. We estimate the interaction between transport and population using a panel of 93 neighbourhoods ("Ortsteile") and 14 periods, approximately every five years from 1870 to 1936 (14 time periods). The interaction between land values and station density can instead be analysed at a level as detailed as 150 x 150 or 300 x 300 metres. Figure A2 illustrates the development of the land values in Reichsmark between 1890 and 1914 on a 300m grid level. Tables A1 to A3 provide the summary statistics for the different samples including the 150m grid sample from the main paper.

Fig. A2. Land values on 300m grid level (in Reichsmark) in 1890 and 1914



1890



1914

Tab. A1. Neighbourhood sample summary statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
$railDens$	overall	0.211	0.3539	0	1.995	N = 1,302
	between		0.268	0.001	0.834	n = 93
	within		0.233	-0.503	1.372	T = 14
$population$	overall	29,743.77	56,296.1	0	354,684	N = 1,302
	between		47,686.39	0	216,328.3	n = 93
	within		30,298.21	-146,107.2	221,449.6	T = 14

Tab. A2. Grid (150m) sample summary statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
$railDens$	overall	0.265	0.276	0.000	1.769	N = 58044
	between		0.217	0.000	0.961	n = 9674
	within		0.171	-0.477	1.233	T = 6
$land\ values$	overall	69.564	151.908	1.000	2180.000	N = 58044
	between		145.259	1.100	1863.000	n = 9674
	within		44.471	-654.716	1219.412	T = 6

Tab. A3. Grid (300m) sample summary statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
$railDens$	overall	0.253	0.269	0.000	1.758	N = 17556
	between		0.214	0.000	0.956	n = 2926
	within		0.162	-0.481	1.208	T = 6
$land\ values$	overall		139.266	1.000	1950.750	N = 17556
	between	62.939	133.872	1.100	1651.143	n = 2926
	within		38.450	-511.376	616.135	T = 6

3.2 Core-periphery sample

Following Levinson (2008), we distinguish between a core region and the periphery as defined by the historian Leyden (1933). The core area serves as a feasible approximation of the area where the vast majority of economic activity took place in historic Berlin. This concentration of commerce in central areas can be rationalised with agglomeration economies that increase firm productivity as discussed in more detail by Ahlfeldt & Wendland (2013) in the context of historic Berlin.

4 Unit root tests

This section provides unit root tests for all models used in the main paper. The test results are illustrated in Table A4. Variables are logarithmised, time-demeaned and Helmert transformed as in the actual analysis. We use the Phillipps-Perron version (PP) of the modified Fisher-type test. The inverse X^2 transformed test statistic (Choi, 2001) rejects the null hypothesis of all panels being non-stationary. Population is rejected at a significance level of 1%. Rail station density turns out to be stationary at a 1% significance level, too. We then apply the Pesaran (2007) unit root test to control for potential cross-section dependence. The test rejects the null hypothesis of the series being non-stationary, I(1), for population and rail station density at a 1% level.

We can only apply the second generation unit root test to the population and rail station density panel used in section 6.4 in this appendix. The other samples are too short and lack a sufficient number of consecutive periods for performing a Pesaran unit root test. However, according to Sarafidis & Robertson (2009) the bias caused by potential cross-section dependence can be reduced when the series are time-demeaned prior to the estimation. We therefore expect the other series to sufficiently fulfil the stationarity requirements, too, as we estimate the PVAR with time-demeaned and forward-mean-differed series. Having confirmed the stationarity of our (transformed) series, we can move on to the actual estimation of the PVAR.

Tab. A4. Panel unit root tests

		Choi (2001) Phillipps-Perron		Pesaran (2007)
Variable		Neighbourhood sample	Grid 150m sample	Neighbourhood sample
pop	test statistic	23.592***	-	-4.552***
	p-value	0.0000	-	0.000
$railDens$	test statistic	51.403***	93.561***	-4.646***
	p-value	0.000	0.000	0.000
LV	test statistic	10.198***	185.959***	-
	p-value	0.000	0.000	-

Notes: (1) Variables shown are logarithmised, time-demeaned and Helmert transformed, (2) *** p<0.01, ** p<0.05, * p<0.1., (3) Pesaran (2007) Test only applicable to sampled used in the bivariate population and station density models (A5.5).

Our baseline transport access measure assumes that the impact of an additional station in a neighbourhood depends on the number of already existing stations, which imposes a strong form of complementarity of the services offered. We replicate our multivariate land use model using the distance to the nearest station, which is a popular measure in the literature and imposes that stations are perfect substitutes.

The results of the panel unit root test for the new transport variable on a neighbourhood level are reported in Table A5. According to the Phillipps-Perron version of the Fisher-type test, the transport measure is stationary at a 1% level.

Tab. A5. Panel unit root tests (distance to nearest station)

Variable		Choi (2001) Phillips-Perron Neighbourhood sample
$distRail$	test statistic	6.521***
	p-value	0.000

Notes: (1) Variables shown are logarithmised, time-demeaned and Helmert transformed, (2) *** p<0.01, ** p<0.05, * p<0.1.

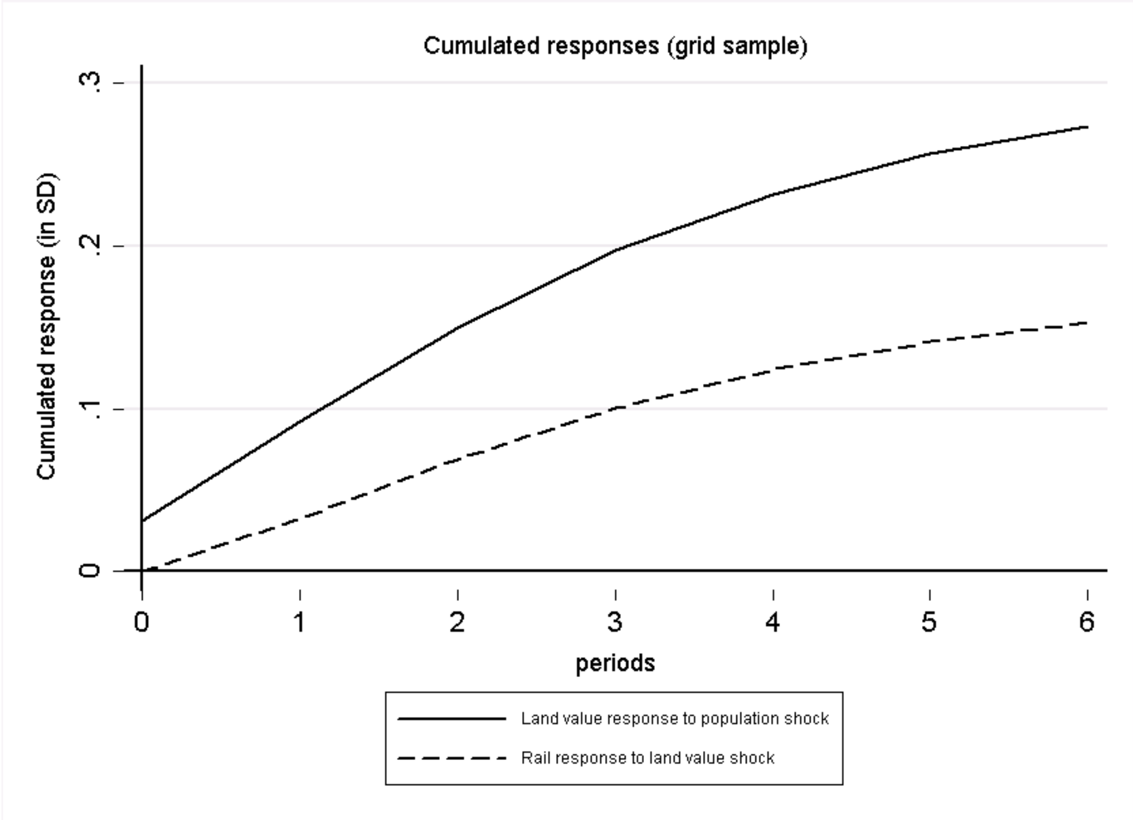
5 Land value: Bivariate demand supply models

This section provides a detailed discussion of the robustness checks briefly summarised at the end of section 3.1 in the main paper.

5.1 Cumulative IRF

The cumulative impulse responses for the bivariate PVAR reported in the main paper (standard IRF in Fig. 4) are shown in Figure A3. The curves illustrate the response to a shock accumulated over all time periods. The solid line indicates the cumulative land value response. The response is immediate and grows at a decreasing rate over time. It is about twice as strong as the planners' cumulated response (dotted line) to a land value shock.

Fig. A3. Cumulative impulse responses: Bivariate land value model



5.2 Modifiable areal unit problem (MAUP)

To address concerns regarding the MAUP, we evaluate the extent to which aggregation to neighbourhoods affects the results in the bivariate land value station density models. We aggregate our data on land values and on station density to a 300m (see above) grid square level.

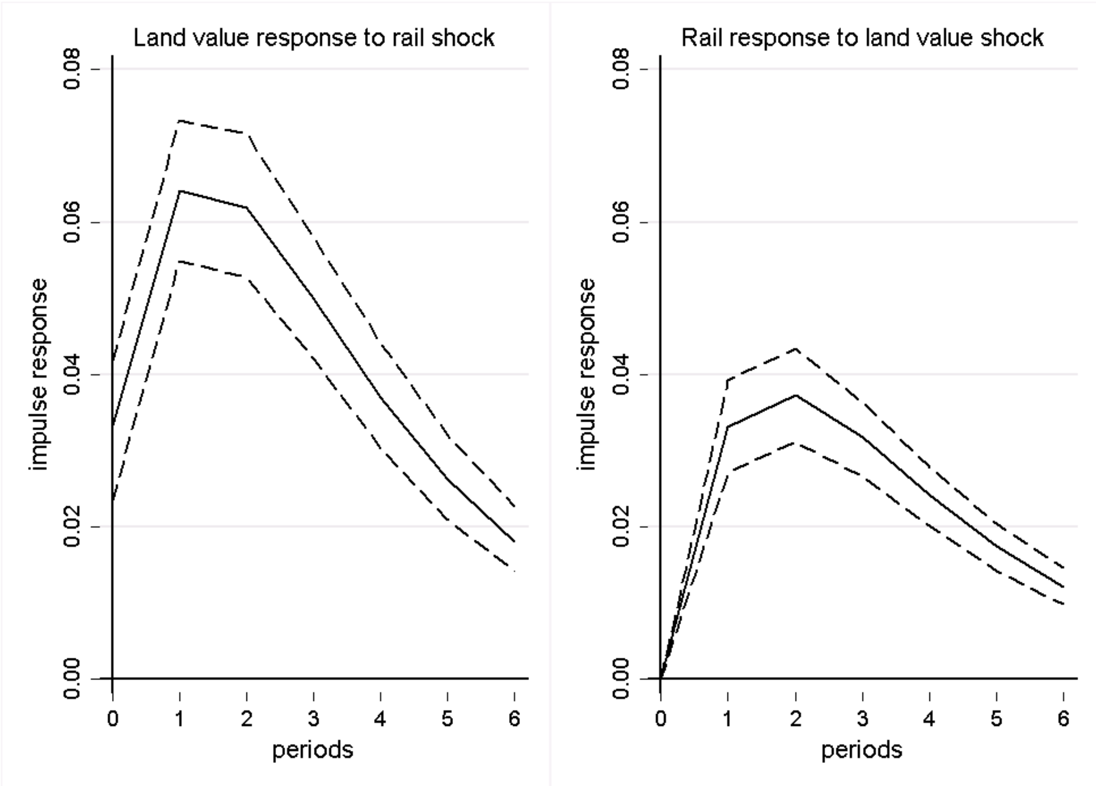
Reduced form results for the 300m grid level bivariate models are reported in Table A6 with their respective IRF in Figure A4. The IRF patterns are very similar to the ones derived from the 150m grid sample used in the main text (Figure 4 in the main text). Overall, the impulse responses are slightly more pronounced in terms of standard deviations.

Tab. A6. Reduced form: Bivariate land value model (300m grid level)

	(1) log land value (t)		(2) log station density (t)	
	Coeff.	S.E.	Coeff.	S.E.
log land value (t-1)	0.606***	(0.014)	0.058***	(0.006)
log station density (t-1)	0.097***	(0.010)	0.519***	(0.028)
Obs.		10,369		

Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Fig. A4. Impulse responses: Bivariate land value model (300m grid level)



Notes: IRF illustrate the effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation. Dashed lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

The similarity of the IRF as derived at different levels of spatial aggregation relieves concerns that a MAUP may be present in the data.

5.3 Transport accessibility measurement

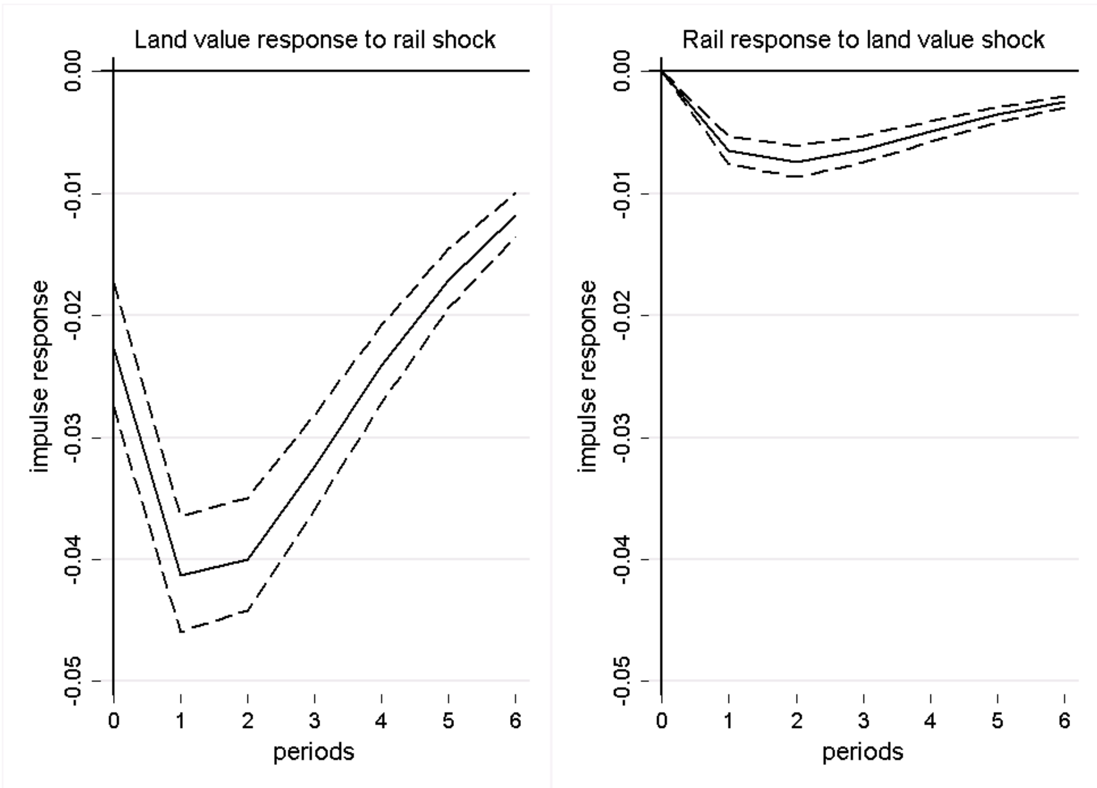
Reduced form estimates for the bivariate 150m grid level Panel VAR are shown in Table A7 with its respective IRF in Figure A5. It is important to recall that the alternative transport measure is a distance measure, i.e. a positive transport shock is now reflected by a reduction in the distance to the nearest station. Hence land values react negatively to an increase distance to station indicating that a transport improvement is capitalised into land values (left panel). Planners also respond to land value shocks, decreasing distance to stations with higher land values. Overall, the alternative transport measure estimates are in line with the results in the main text.

Tab. A7. Reduced form: Bivariate land value model (150m grid level), nearest station

	(1) log land value (t)		(2) log nearest station (t)	
	Coeff.	S.E.	Coeff.	S.E.
log land value (t-1)	0.616***	(0.007)	-0.011***	(0.011)
log nearest station (t-1)	-0.127***	(0.011)	0.531***	(0.013)
Obs.		38,696		

Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Fig. A5. Impulse responses: Bivariate land value model (150m grid level), nearest station



Notes: IRF illustrate the effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation. Dashed lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

5.4 Announcement network

According to the weak-form efficient market hypothesis, asset prices incorporate all available information that is publicly available. We would therefore expect transport infrastructure to be capitalised into prices at the time of announcement (not completion). To accommodate such effects we reran our benchmark models using historical station density measures based on the dates of announcements.[1]

The reduced form results of the bivariate (announced) station density model are given in Table A8, its respective IRF are shown in Figure A6. As stated by the weak-form efficient market hypothesis, land values respond immediately to the announcement of new stations (left graph). We observe a slightly more spikey pattern which is, however, very similar in terms of magnitudes compared to the benchmark model discussed in the main text. In line with the actual station density, the IRF shown in the right graph of Figure A6 stresses the empiric relevance of the reverse impact of economic output (as capitalised in land value) on the planning process. The transport response is somehow weaker in terms of standard deviation than the land value response to transport innovation. This is also in line with the bivariate model in the main text.

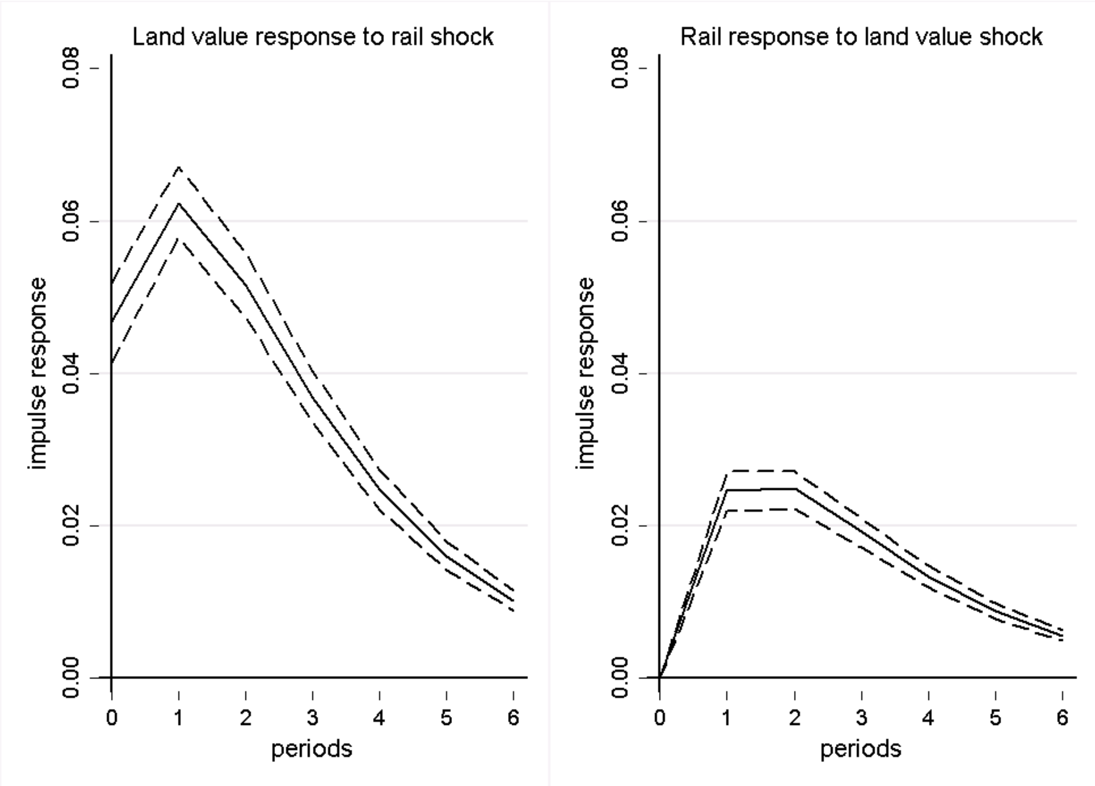
Tab. A8. Reduced form: Bivariate land value model (150m grid level), announced network

	(1) log land value (t)		(2) log station density (t)	
	Coeff.	S.E.	Coeff.	S.E.
log land value (t-1)	0.594***	(0.008)	0.040***	(0.003)
log station density (t-1)	0.086***	(0.006)	0.417***	(0.013)
Obs.		34,046		

Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

[1] The definition of the announcement dates is based on Dudczak & Dudczak (2012), Kurpjuweit & Meyer-Kronthaler (2009), Mauruszat (2011), Loop (2007), Luisenstädtischer Bildungsverein e.V. (2012), Senatsverwaltung für Stadtentwicklung und Umwelt (2012) and Straschewski (2011). Where no information on announcement is available we define the beginning of construction works as the announcement.

Fig. A6. Impulse responses: Bivariate land value model (150m grid level), announced network



Notes: IRF illustrate the effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation. Dashed lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

6 Land use: Multivariate demand supply models

This section provides a detailed discussion of the robustness checks briefly summarised at the end of section 3.2 in the main paper.

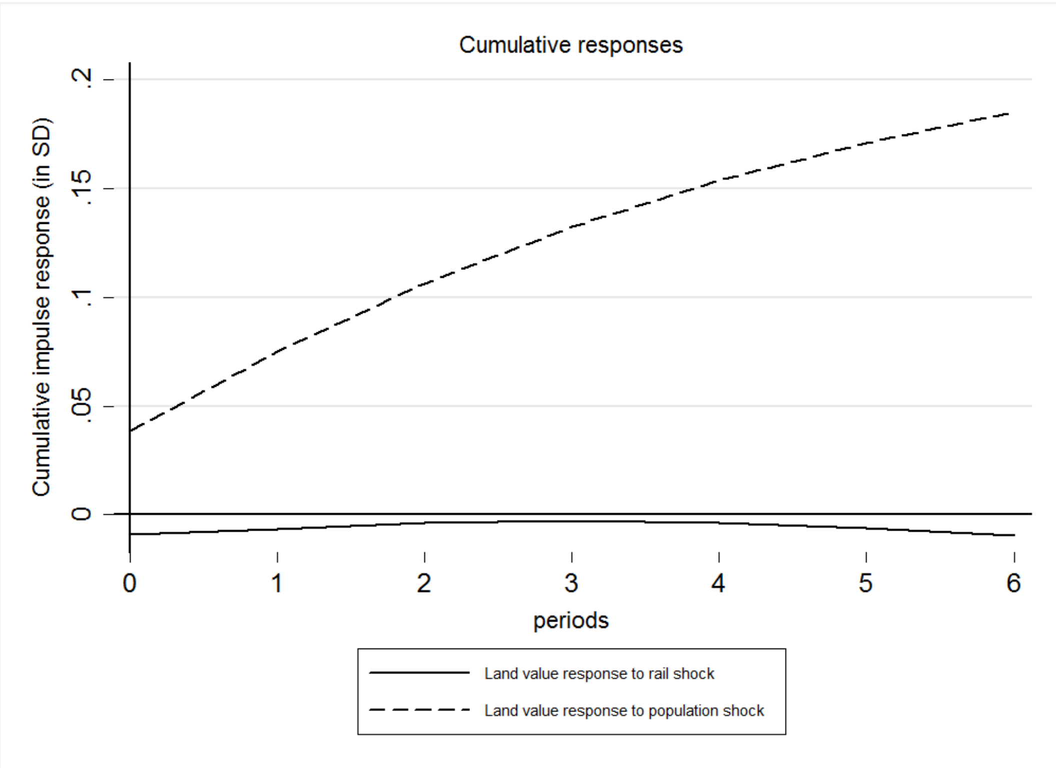
6.1 Cumulative IRF

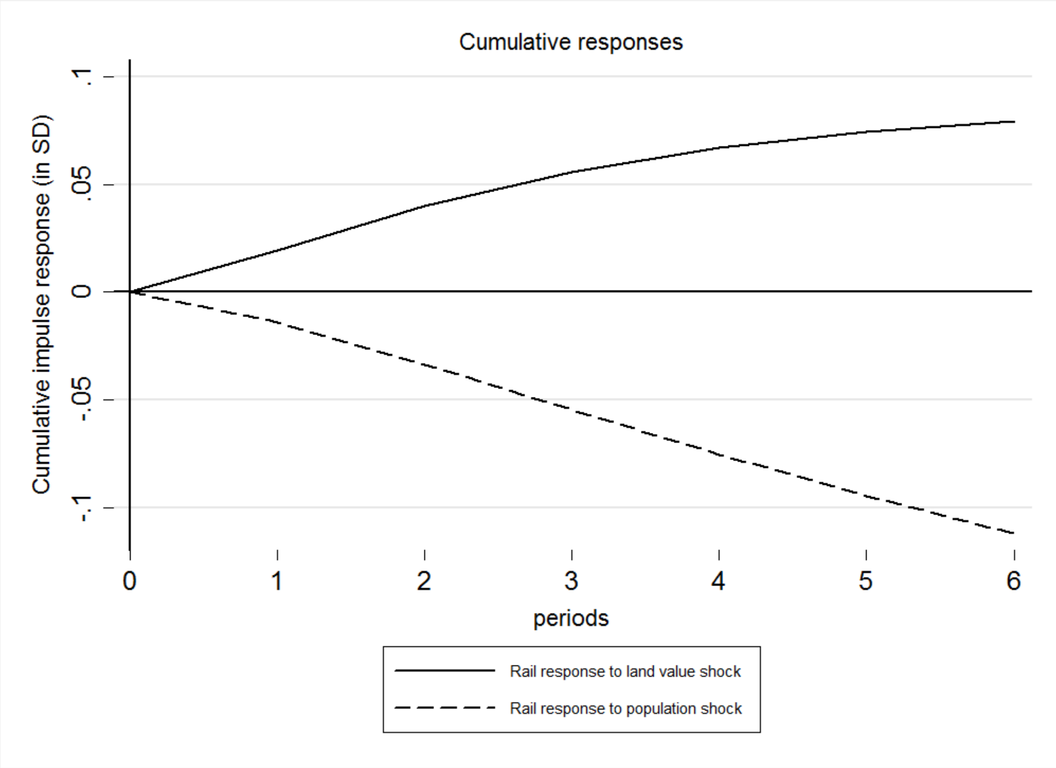
The cumulative impulse responses for the multivariate demand supply land use model (Fig. 5 in the main paper) are shown in Figure A7. We begin with the cumulative response of land values to transport as well as to population shocks (upper bar). As indicated in the main part, the transport innovation induces an immediate land value adjustment. We have argued that the contemporaneous effect reflects a short-run adjustment (increase in demand with a constant building stock). Over time the effect diminishes due to an adjustment in the intensity of land use. Land values steadily rise in response to population shocks over time. This is explained by a relatively slow adjustment in building stock per land unit which hinders a quick adjustment to the new equilibrium land value.

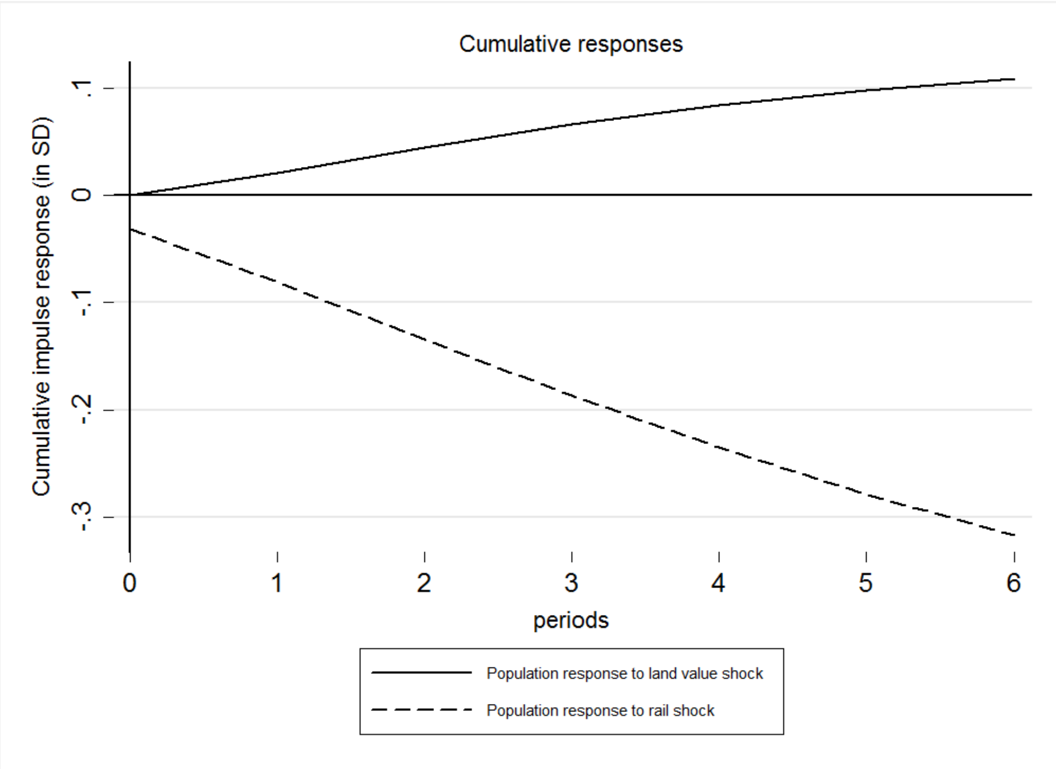
The cumulative transport response is shown in the middle panel of Figure A7. Rail responds positively to land values over time and negatively to population. We expect a negative population shock to be an indicator of a positive shock in commercial land use, especially in the short run where the supply of floor space is highly inelastic. The planner's negative response to population is therefore interpreted as an increase in transport supply owing to a substitution of residential for commercial space. Moreover, the positive transport reaction to land values indicates that there is a substitution of land uses and no general negative (economic) shock to the city.

Population declines negatively with transport and positively but insignificantly with land value shocks over time. This strengthens our interpretation of transport innovation leading to a relative increase in the intensity of commercial use. Limits to densification result in a displacement of residents.

Fig. A7. Cumulative impulse responses: Multivariate land use model (neighbourhood level)







Notes: Cumulative IRF illustrate accumulated effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation. In the legend, the response variable is written outside the bracket while the shock is inside the parentheses.

6.2 Transport accessibility measurement

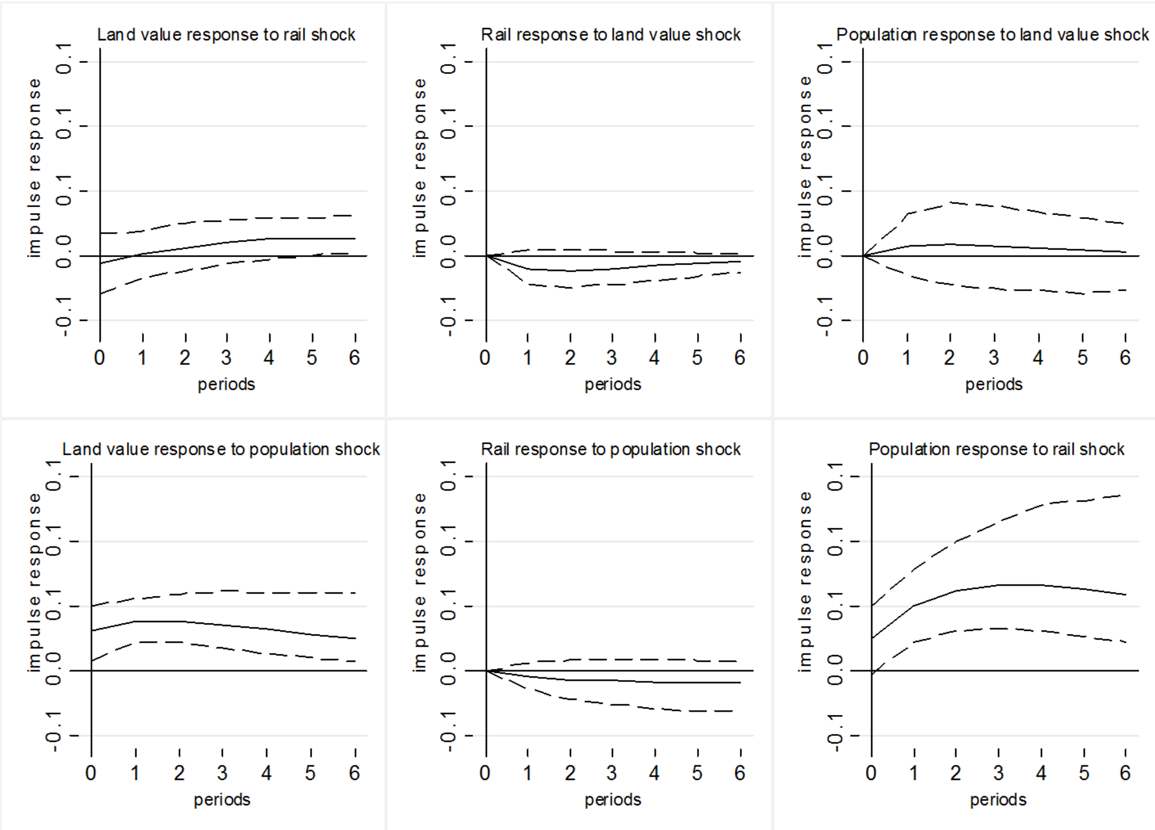
The reduced form estimates of the multivariate land use model using an alternative transport measure are shown in Table A9 and, as the IRF depicted in Figure A8, are comparable in terms of quality and quantity with the results in the main text. Due to the reverse interpretation of the transport measure, the majority of the IRF look like a mirrored version of the IRF from Figure 5 in the main text. The only notable difference to the benchmark model reported in the main paper is that the transport supply responses to land value (upper middle) and population shocks (bottom middle) are not statistically significant.

Tab. A9. Reduced form: Multivariate land use model (neighbourhood level), nearest station

	(1) log land value (t)		(2) log nearest station (t)		(3) log population (t)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
log land value (t-1)	0.182***	(0.038)	-1E-5	(0.004)	-0.126***	(0.061)
log nearest station (t-1)	-0.327	(0.346)	0.782***	(0.060)	2.589***	(0.524)
log population (t-1)	0.109***	(0.032)	-0.015***	(0.004)	0.972***	(0.073)
Obs.			169			

Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Fig. A8. Impulse responses: Multivariate land use model (neighbourhood level), nearest station



Notes: IRF illustrate the effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation. Dashed lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

6.3 Announcement network

Estimates for the multivariate models using the announced network are shown in Table A10 (reduced form results) and in Figure A9 (IRF). The qualitative interpretations remain remarkably similar to the benchmark models. The main difference is the planners' response to population shocks, which is insignificant (bottom middle). Forward-looking planners did not incorporate residential developments when planning new transport routes. This is in line with the idea that planners are more likely to follow commercial activity.

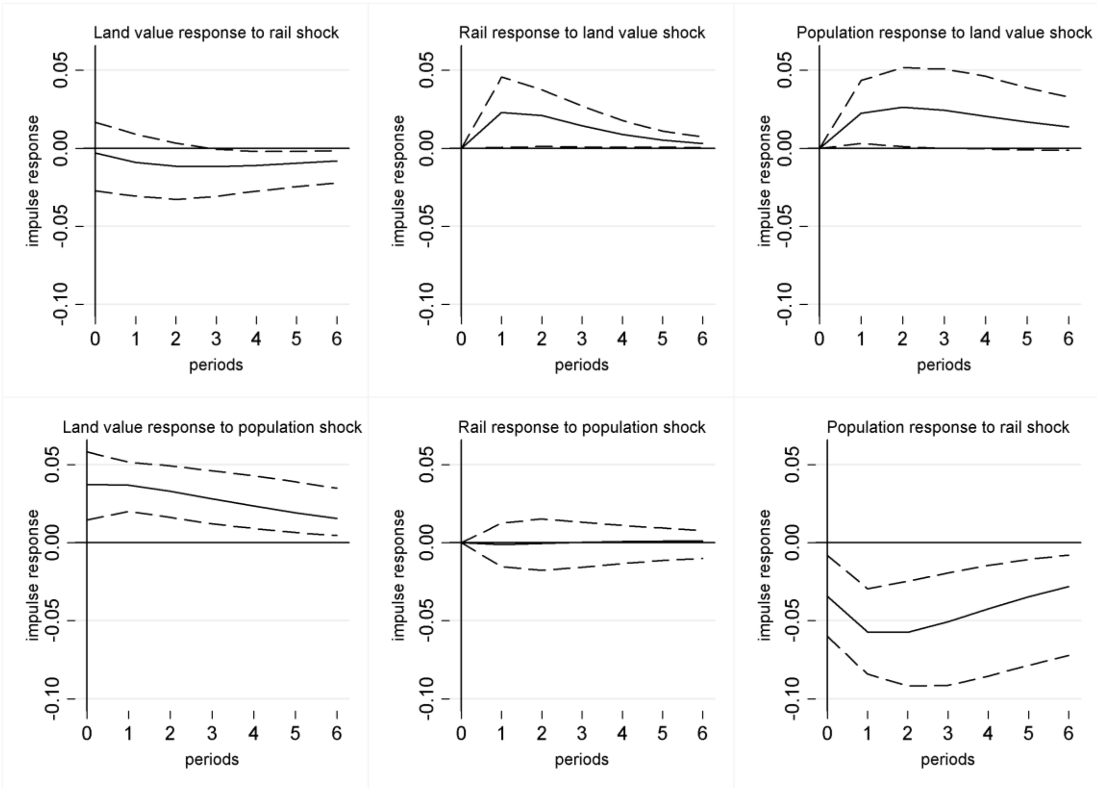
Summing up, we find evidence of the weak-form efficient market hypothesis. Using an announced network instead of the actual one does not significantly change our findings on the interaction between transport and land values/population.

Tab. A10. Reduced form: Three variable PVAR model (neighbourhood level), announced network

	(1) log land value (t)		(2) log station density (t)		(3) log population (t)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
log land value (t-1)	0.525***	(0.074)	0.124*	(0.075)	0.121**	(0.067)
log station density (t-1)	-0.016	(0.033)	0.415***	(0.124)	-0.109***	(0.064)
Log population (t-1)	0.086***	(0.039)	-0.028	(0.048)	0.765***	(0.087)
Obs.			188			

Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Fig. A9. Impulse responses: Three variable PVAR model (neighbourhood level), announced network



Notes: IRF illustrate the effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation. Dashed lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

6.4 Land use: Bivariate supply demand model

We now introduce an alternative bivariate model which uses population density instead of land values and hence follows the originally discussed Levinson (2008) approach. As discussed in detail in section 2 of the main paper, the analysis then needs to be disaggregated by land use in this model. The reason is that depending on the relative attractiveness of an area to firms and residents, population will be either attracted to an area with improved transport supply, or displaced by firms that relocate to the area.

We firstly report the results for the same sample we use for the benchmark multivariate model. Secondly, we make use of the full sample and extend the number of periods up to 14.

The reduced form estimates based on the sample used in the multivariate model are reported in Table A12 for the total, Table A13 for the core and Table A14 for the periphery sample. The IRF are illustrated in Figure A10 (total sample) and Figure A11 (core/periphery). Adopting the recursive order assumptions, we assume that population responds to contemporaneous transport shocks while transport only responds to lagged population and transport innovations.

We begin the brief analysis with the full multivariate model sample. The immediate population response is positive but insignificant in the following periods (Figure A10, left). As already indicated by the multivariate benchmark model there is no (significant) transport response with respect to population (right). Moving on to the spatial subsample IRF depicted in Figure A11, population response is negative in the core (top-left) and, except from a positive contemporary response, insignificant in the periphery (bottom-left). These patterns are in line with our interpretation of residents being displaced by firms following a positive transport shock in central areas. Also the reverse planner response with respect to population is in line with previous findings. Planners respond negatively to population shocks in the periphery (top-right). We interpret this as a positive response to an increase in the intensity of commercial use. We do not observe a significant transport reaction to population shocks in the periphery either (bottom-right).

Tab. A11. Reduced form: Population station density model (neighbourhood level), total sample, multivariate model sample

	(1) log population (t)		(2) log station density (t)	
	Coeff.	S.E.	Coeff.	S.E.
log population (t-1)	0.893***	(0.028)	0.033	(0.040)
log station density (t-1)	-0.010	(0.021)	0.705***	(0.127)
Obs.		559		

Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Tab. A12. Reduced form: Population station density model (neighbourhood level), core sample, multivariate model sample

	(1) log population (t)		(2) log station density (t)	
	Coeff.	S.E.	Coeff.	S.E.
log population (t-1)	0.531***	(0.151)	-0.244***	(0.082)
log station density (t-1)	-0.308***	(0.099)	0.780***	(0.113)
Obs.		77		

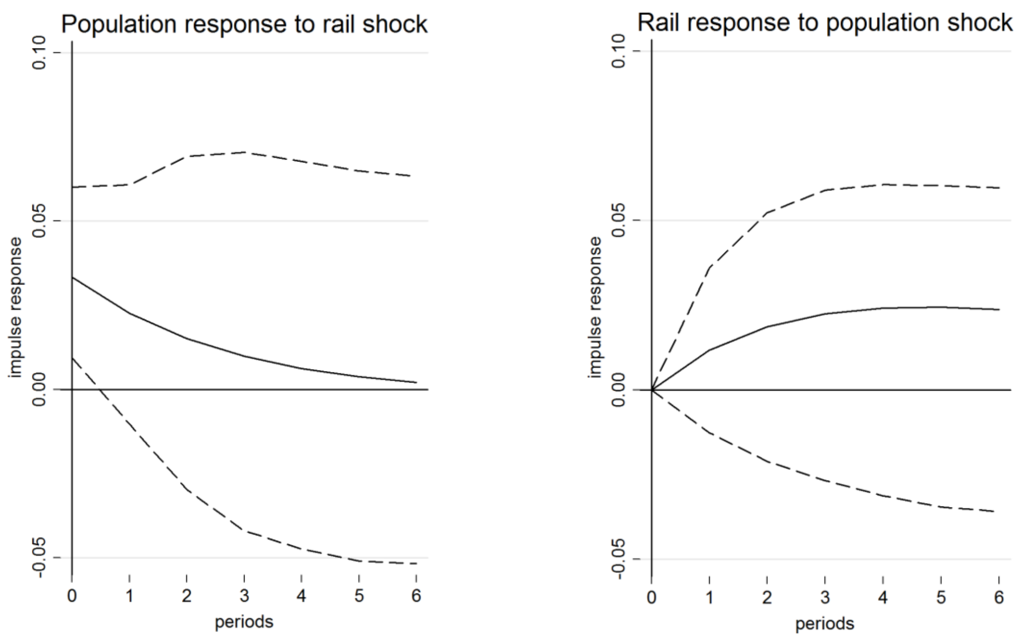
Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Tab. A13. Reduced form: Population station density model (neighbourhood level), periphery sample, multivariate model sample

	(1) log population (t)		(2) log station density (t)	
	Coeff.	S.E.	Coeff.	S.E.
log population (t-1)	0.889***	(0.029)	0.045	(0.045)
log station density (t-1)	-0.007	(0.021)	0.700***	(0.130)
Obs.		482		

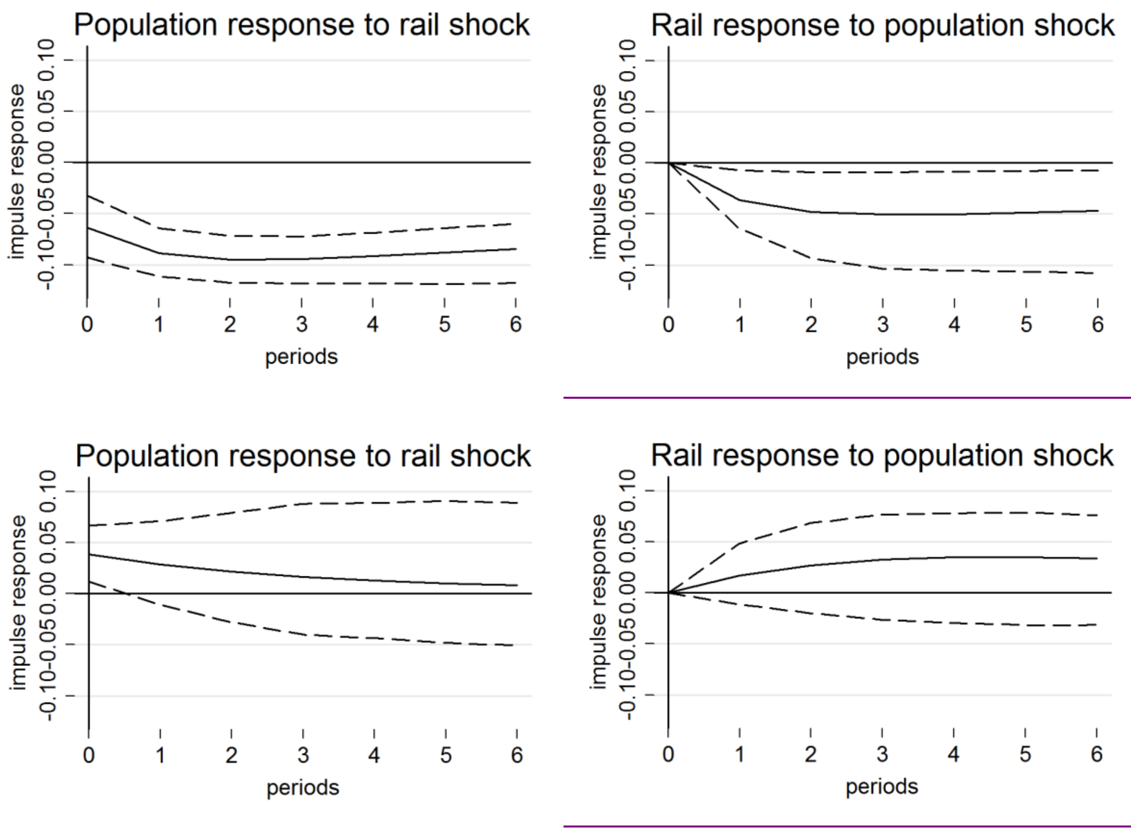
Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Fig. A10. Impulse responses: Population station density model (neighbourhood level), multivariate model sample



Notes: IRF illustrate the effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation. Dashed lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

Fig. A11. Impulse responses: Population station density model (neighbourhood level), core (upper graphs) and periphery (lower graphs), multivariate model sample



Notes: IRF illustrate the effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation. Dashed lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

Table A14 (total sample), Table A15 (core) and Table A16 (periphery) provide the reduced form estimates for the bivariate population model making use of all available time periods. The respective IRF are shown in Figure A12 (total sample) and Figure A13 (core/periphery). The patterns are comparable to the previously reported ones and further strengthen our interpretations. Again, population responds negatively to transport improvements in the core (Figure A13, top-left). In the periphery (Figure A13, bottom-left) and for the total sample (Figure A12, left) the response is positive instead. The displacement of residents is therefore not restricted to the observation period used in the main analysis but becomes even clearer when extending the panel. Transport supply responses are found to be insignificant in all spatial samples, indicating that land use changes might have been particularly influential in determining transport improvements during the period we have focussed on in our benchmark models (1890–1915).

Tab. A14. Reduced form: Population station density model (neighbourhood level), total sample, full sample

	(1) log population (t)		(2) log station density (t)	
	Coeff.	S.E.	Coeff.	S.E.
log population (t-1)	0.800***	(0.030)	-0.029	(0.031)
log station density (t-1)	0.029***	(0.015)	0.745***	(0.079)
Obs.	1,015			

Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Tab. A15. Reduced form: Population station density model (neighbourhood level), core sample, full sample

	(1) log population (t)		(2) log station density (t)	
	Coeff.	S.E.	Coeff.	S.E.
log population (t-1)	0.818***	(0.049)	-0.139	(0.118)
log station density (t-1)	-0.086***	(0.020)	0. 803***	(0.072)
Obs.	132			

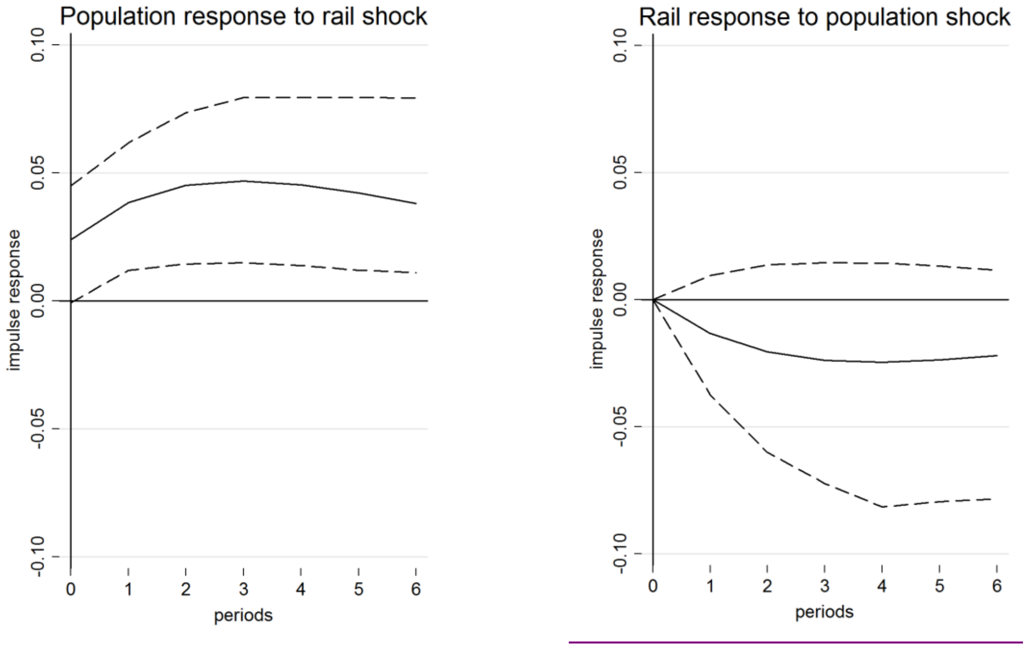
Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Tab. A16. Reduced form: Population station density model (neighbourhood level), periphery sample, full sample

	(1) log population (t)		(2) log station density (t)	
	Coeff.	S.E.	Coeff.	S.E.
log population (t-1)	0.792***	(0.031)	-0.022	(0.035)
log station density (t-1)	0.036***	(0.016)	0.739***	(0.084)
Obs.	883			

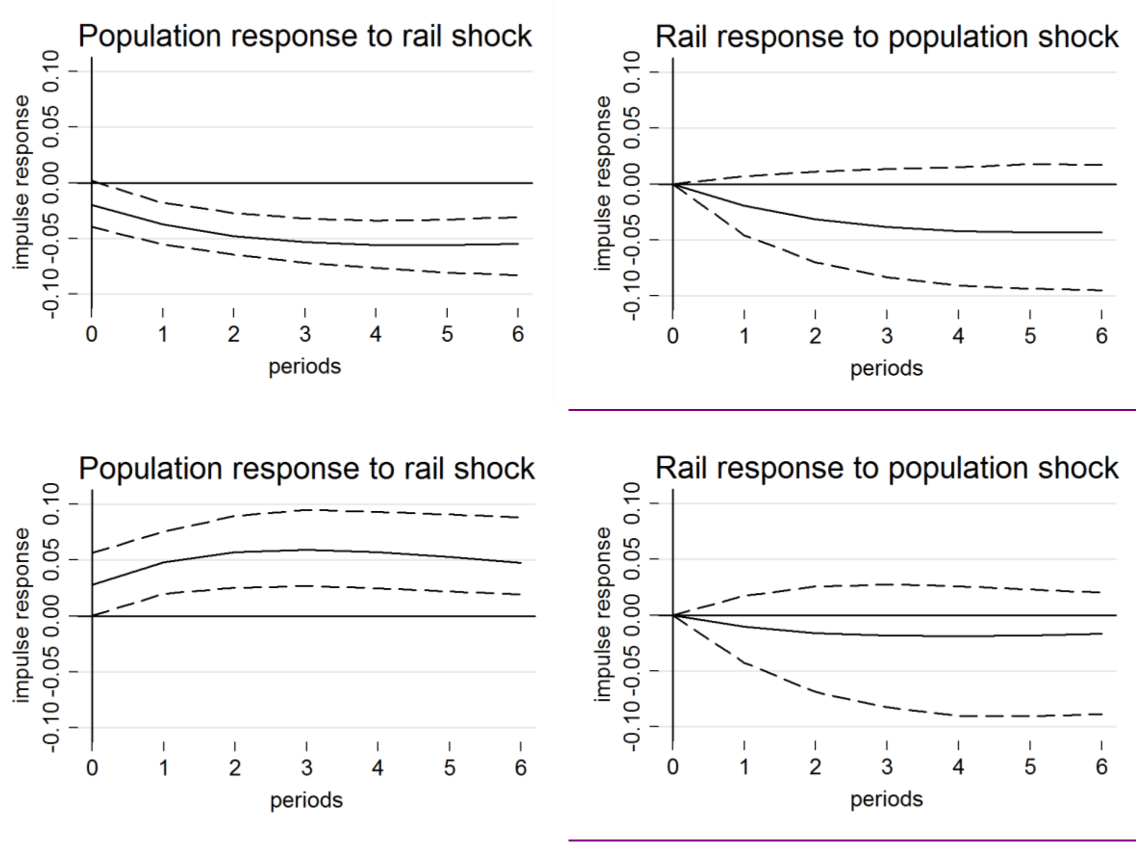
Notes: 1-lag VAR is estimated by GMM. All variables are in logs time-demeaned and Helmert transformed. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Fig. A12. Impulse responses: Population station density model (neighbourhood level), full sample



Notes: IRF illustrate the effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation. Dashed lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

Fig. A13. Impulse responses: Population station density model (neighbourhood level), core (upper graphs) and periphery (lower graphs), full sample



Notes: IRF illustrate the effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation. Dashed lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

Summing up our results, we find that the relationship between population and rail density is qualitatively distinct in the core and periphery sample. While new infrastructure tends to displace residents in the core area, it attracts residents in the periphery area.

7 Complementary IV analyses

In this section we present a complementary analysis of the causal effect of transit supply on land value and land using a panel IV strategy. A quantitative comparison of the results to the PVAR model results is provided in section 3.3 of the main paper. We begin with the impact of the station density of land value in section 7.1 before we discuss the impact of station density on land use in section 7.2.

7.1 Land value

This section presents the empirical strategy and the results of a complementary analysis of the impact of station density on land value. Our approach shares similarities with the research design employed by Gibbons & Machin (2005) and Ahlfeldt & Wendland (2009). To these established approaches we add an IV, which is supposed to restrict the variation in station density used for identification to the fraction that is presumably attributable to exogenous planning objectives.

7.1.1 Strategy

One advantage of our empirical setting is that we can exploit substantial variation not only across space, but also over a number of consecutive periods. We can therefore allow for a large degree of unobserved spatial heterogeneity in levels and trends in our identification strategy to strengthen the causal inference. Our point of departure is an empirical model that describes the (log) land value as a function of station density T. We allow for a fixed composite location amenity L, which impacts on the level of land value and the long-run yearly trend (t), as well as trends that are specific to a period (Y) and neighbourhood (O).

$$\log V_{it} = \alpha_1 \log T_{it} + \sum_i \beth_i L_i \times t + \sum_i \omega_i L_i + \sum_j \sum_t \aleph_{ij} \left(O_j \times Y_t \times t\right) + \varepsilon_{it}$$

We remove the unobserved location-specific effect on levels at the plot level by taking first differences (Gibbons & Machin, 2005). Our final empirical specification then describes the relationship between changes in (log) land values and (log) station density conditional on a set of plot fixed effect and neighbourhood × period effects.

$$\Delta \log V_{it} = \alpha_1 \Delta \log T_{it} + \sum_i \beth_i L_i + \sum_j \sum_t \aleph_{ij} \left(O_j \times Y_t\right) + \Delta \varepsilon_{it}$$

With this specification we only identify variation within neighbourhoods in a given year. We follow Gibbons et al (2012) in arguing that at a very small spatial level the variation provided by the infrastructure is as good as random because the exact routing is determined by local particularities that are exogenous to economic development (e.g. soil conditions and other geographical features). Our neighbourhoods are relatively small areas; frequently smaller than 10 sq km.[2] Under the identifying assumption that the routing

[2] For comparison, Gibbons, et al (2012) use UK wards as a unit of analysis with an average size of about 16.6 sq km.

within a neighbourhood is exogenous to demand OLS estimation of this model will deliver unbiased estimates of the causal effect of station density on land prices.

To further strengthen identification and to relax this identifying assumption we make use of an IV based on a counterfactual network used by Ahlfeldt & Wendland (2011). The network consists of straight lines that connect the CBD to the most important nearby towns as well as an emerging secondary centre (the Kurfürstendamm). We distributed counterfactual stations every 1,089 metres along the IV tracks before computing a counterfactual density measure. The average distance between railway stations in 1915 was 1,089 metres. By definition, this is a time-invariant instrument. To introduce time variation into our IV we run $i{\times}t$ locally weighted regressions (LWR) (Cleveland & Devlin, 1988; McMillen, 1996) of station density on counterfactual station density for each year and every plot. In each iteration we weight all observations within a given year based on the distance to a plot i using the following kernel function: $w_{i\bar{i}} = \exp(-\tau^2 D_{i\bar{i}}^2)$, where we set τ so that the function flattens out after about 5km. The predicted values of these regressions form a panel variable that varies over space and time, provides a reasonable fit to the overall evolution of the city structure, and yet removes some of the local co-variation between land values and actual station density. The variable qualifies as an instrument for station density because, by construction, it has predictive power, and because conditional on being located within a certain neighbourhood in a certain time period, the variation provided by the counterfactual network is as good as random.

To rationalise this strategy we argue that being closer (i.e. in a denser area with respect to IV rail stations) to the potential transport corridors defined by the counterfactual network increases the chance of being connected to the network. At the same time, being closer to the hypothetical network conditional on distance to the CBD (and other amenities) is as good as random. Put simply, our IV restricts the variation in changes of station density to the portion that can reasonably be assumed to be exogenous.

7.1.2 Results

The results of the complementary IV analysis are presented in Table A17. We begin with the baseline specification (column 1), where we regress the change in land values on the change in station density (both in logs) while controlling for period fixed effects. We find a positive relation significant at the 1% level. Controlling for individual heterogeneity (column 2) as well as period fixed effect interacted with neighbourhood effects (column 3) slightly increases the coefficient. Our instrumental variable is introduced from column (4) onwards; transport continues to positively and significantly drive land values throughout all remaining specifications. The results consistently point to an elasticity of about 12%. Our preferred estimate implies that a one SD increase in station density leads to a 0.08 SD in land value. The IV estimates are within the same range, suggesting that the OLS models are not biased due to reverse causality,

Tab. A17. Complementary IV analyses 1: Land values

	(1)	(2)	(3)	(4)	(5)	(6)
				Δ log land values		
Δ log station density	0.110***	0.145***	0.124***	0.111***	0.153***	0.128***
	(0.010)	(0.011)	(0.012)	(0.009)	(0.010)	(0.011)
Plot FE	NO	YES	YES	NO	YES	YES
Period FE	YES	YES	NO	YES	YES	YES
Ortsteil x period FE	NO	NO	YES	NO	NO	YES
IV	NO	NO	NO	YES	YES	YES
N	48370	48370	48370	48370	48370	48370
F (first stage)	-	-	-	660.62	20862.23	636.63

Notes: Instrument variable: Log station density of counterfactual network. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

The complementary land value analysis builds on the popular instrumental variable approach. By repeating a similar exercise as in the bivariate Panel VAR model but using a distinct and well-established empirical approach, we try to add some validity to the PVAR estimates discussed in the main paper. The consistency of the findings is reassuring given the fundamental differences of the models and the underlying data transformations. The (dynamic) panel model is not estimated simultaneously and describes the relationship between variables, while the PVAR approach describes the relation between innovations (shock view).

7.2 Land use

In this section we present causal estimates of changes in station density on land use using a strategy that is similar to the one presented in the previous section.

7.2.1 Land use data and empirical strategy

The information on land use was extracted from a series of map publications that provide detailed reconstruction of real land use in 1880, 1910 and 1940 (Aust, 1986, 1987). On these maps, each parcel of land is assigned to one of the following categories: industrial, public, residential, business or mixed use. From the maps it is also evident if a parcel at a given point in time was undeveloped. Lastly, the maps show the boundaries of green space, water spaces and overground rail tracks.

To process this information in a statistical analysis, we intersect the raster data with our 150m grid, whose cells form the cross-sectional unit of a panel data set. This approach flexibly accommodates changes in land use at a fine level without imposing arbitrary official boundaries. As such, the grid does not imply a density bias, i.e. implicitly higher weights to smaller geographic areas typically encountered in more central areas. We calculate the individual share of land use each grid cell covers. The empirical strategy we use to estimate the causal effect of increasing the supply of transport services on land use is similar to the one discussed in the previous section. We replicate the land value in perfect analogy but replace the dependent variable with the log share of land within a 150m grid cell that is at least partially used for commercial purposes (commercial land and mixed use).

7.2.2 Results

Estimation results are shown in Table A18. We use the same set of specifications as before. The results confirm the implications of the PVAR model in that an increase in station density significantly increases the share of commercial land use. In our preferred IV models (columns 5 and 6) the elasticity is about 0.3.

As a further interpretation from the PVAR models we derive that the displacement of residential for commercial use occurs at a faster rate within the urban core area. To accommodate the treatment heterogeneity, we allow for an interaction effect between station accessibility and distance to the CBD. The results (Table A19) reported below are supportive of this heterogeneity, especially in our preferred IV models (columns 5 and 6).

Tab. A18. Complementary IV analyses 2: Land use

	(1)	(2)	(3)	(4)	(5)	(6)
	Δ log share of land with significant commercial use					
Δ log station density	0.267***	0.246***	0.199***	0.359***	0.310***	0.303***
	(0.015)	(0.018)	(0.029)	(0.015)	(0.019)	(0.041)
Plot FE	NO	YES	YES	NO	YES	YES
Period FE	YES	YES	NO	YES	YES	YES
Ortsteil x period FE	NO	NO	YES	NO	NO	YES
IV	NO	NO	NO	YES	YES	YES

N	8914	8914	8914	8914	8914	8914
F (first stage)	-	-	-	305.68	8413.72	655.51

Notes: Instrument variable: Log station density of counterfactual network. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Beta coefficients in squared bracket.

Tab. A19. Complementary IV analyses 2: Land use and treatment heterogeneity

	(1)	(2)	(3)	(4)	(5)	(6)
	Δ log share of land with significant commercial use					
Δ log station density	0.316***	0.420***	0.430***	0.318***	0.433***	0.638***
	(0.023)	(0.038)	(0.064)	(0.023)	(0.043)	(0.090)
Δ log station density	-0.012***	-0.043***	-0.056***	0.011**	-0.031***	-0.089***
x distance to CBD	(0.004)	(0.007)	(0.012)	(0.005)	(0.010)	(0.022)
Plot FE	YES	YES	YES	YES	YES	YES
Period FE	YES	YES	NO	YES	YES	YES
Ortsteil x period FE	NO	NO	YES	NO	NO	YES
IV	NO	NO	NO	YES	YES	YES
N	8914	8914	8914	8914	8914	8914
F (first stage)	-	-	-	501.28	3622.59	204.43

Notes: Instrument variable: Log station density of counterfactual network. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Beta coefficients in squared bracket.

References

Ahlfeldt, G. M., & Wendland, N. (2009). Looming Stations: Valuing Transport Innovations in Historical Context *Economics Letters, 105*(1), 97-99. doi: http://dx.doi.org/10.1016/j.econlet.2009.06.010

Ahlfeldt, G. M., & Wendland, N. (2011). Fifty years of urban accessibility: The impact of the urban railway network on the land gradient in Berlin 1890-1936. *Regional Science and Urban Economics, 41*(2), 77-88.

Ahlfeldt, G. M., & Wendland, N. (2013). How polycentric is a monocentric city? Centers, spillovers and hysteresis. *Journal of Economic Geography, 13*(1), 53-83. doi: 10.1093/jeg/lbs013

Aust, B. (Cartographer). (1986). Die städtebauliche Entwicklung Berlins - Berlin um 1910.

Aust, B. (Cartographer). (1987). Die städtebauliche Entwicklung Berlins - Berlin um 1880.

Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance, 20*, 249-272.

Cleveland, W. S., & Devlin, S. J. (1988). Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association, 83*(403), 596-610.

Dudczak, J., & Dudczak, U. (2012). Berliner-Bahnen.de – Bahnen im Berliner Raum Retrieved August 14th, 2012, from http://www.berliner-bahnen.de

Elkins, T., & Hofmeister, B. (1988). *Berlin - The spatial structure of a divided city*. London/New York: Mehuen.

Gibbons, S., Lyytikäinen, T., Overman, H., & Sanchis-Guarner, R. (2012). New Road Infrastructure: the Effects on Firms. *SERC Discussion Paper 117*.

Gibbons, S., & Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics, 57*(1), 148-169.

Gottwaldt, A. (1994). *Das Berliner U- und S-Bahnnetz. Eine Geschichte in Streckenplänen*. Berlin.

Kiebert, W. (2004). *Die Berliner S-Bahn. 1924 bis heute. Geschichte, Technik, Hintergründe*. Stuttgart.

Kiebert, W. (2008). *Die Berliner Nordsüd-S-Bahn*. Stuttgart.

Klünner, H.-W. (1985). *S- und U-Bahnarchitektur in Berlin*. Berlin.

Kurpjuweit, K., & Meyer-Kronthaler, J. (2009). *Berliner U-Bahn. In Fahrt seit hundert Jahren* (2 ed.). Berlin.

Levinson, D. (2008). Density and dispersion: the co-development of land use and rail in London. *Journal of Economic Geography, 8*(1), 55-77.

Leyden, F. (1933). *Gross-Berlin. Geographie der Weltstadt*. Breslau: Ferdinand Hirt.

Loop, A. (2007). Untergrundbahn Retrieved August 14th, 2012, from http://www.untergrundbahn.de

Luisenstädtischer Bildungsverein e.V. (2012). Berlin von A bis Z Retrieved August 14th, 2012, from http://www.luise-berlin.de

Mauruszat, A. (2011). Das Berliner U-Bahn-Archiv: U-Bahn-Geschichte online seit 2004 Retrieved August 14th, 2012

McMillen, D. P. (1996). One Hundred Fifty Years of Land Values in Chicago: A Nonparametric Approach. *Journal of Urban Economics, 40*(1), 100-124.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics, 22*(2), 265-312.

Sarafidis, V., & Robertson, D. (2009). On the impact of error cross-sectional dependence in short dynamic panel estimation. *Econometrics Journal, 12*(1), 62-81.

Senatsverwaltung für Stadtentwicklung und Umwelt. (2012). Denkmalliste, Denkmalkarte, Denkmaldatenbank Retrieved August 14th, 2012, from http://www.stadtentwicklung.berlin.de/denkmal/liste_karte_datenbank/index.shtml

Straschewski, M. (2011). Geschichte und Geschichten rund um die Berliner S-Bahn Retrieved February 25th, 2011, from http://www.stadtschnellbahn-berlin.de