Public Good Provision and Local Employment – Evidence from Grammar School Closures in East Germany*

Ronny Freier† Martin Simmler‡ Christian Wittrock§

Abstract: This paper assesses the impact of public good provision on the spatial distribution of employment as predicted by a local labor market model that allows for commuting. Using local grammar school closures in East Germany after 2000 in a difference-in-differences estimation framework coupled with an entropy balancing strategy, we find that the school closures triggered a decline in the number of (employed) residents by 11%. The number of local employees, in contrast, decreases by just 5% as in-commuters do not respond to the school closures. Moreover, we find – as predicted – that the school closures reduce house prices and that the share of land used for housing declines, albeit only to a small extent.

Keywords: local schools, employment, house price, land use, commuting.

JEL Classification: J4, R3, H4.

---

*We are grateful to our Editor Stephen Ross and two anonymous referees and Michael Devereux, Ben Lockwood, Nadine Riedel, Wolfram Richter and Eddy Tam as well as participants of the Congress of the German Economic Association in Vienna and the 'Regional Economics and Local Political Economy' Workshop at the University of Siegen for helpful comments and suggestions. We gratefully acknowledge financial support from the German Research Foundation (SI 2050/1-1 and RI 2491/2-1).

†DIW Berlin, Mohrenstrasse 58, 10117 Berlin, Germany, and Technische Hochschule Wildau, Wildau, Germany. E-mail: rfreier@diw.de.

‡Corresponding author: University of Oxford Centre for Business Taxation, Park End Street, OX1 1HP, Oxford, UK, and DIW Berlin. Tel: +44 1865 614845. E-mail: martin.simmler@sbs.ox.ac.uk.

§Ruhr Graduate School in Economics, U Bochum and TU Dortmund. E-mail: christian.wittrock@rub.de

1 Introduction

How important are local public goods such as schools for the spatial distribution of employment? The standard answer to this question is that it depends on the housing supply elasticity. If the housing supply is fixed, inter-jurisdictional differences in public good provision largely clear through house price adjustments. If the housing supply is elastic, the housing stock changes, and the impact on population and employment will be stronger. While the housing supply elasticity is thus important, the 'commuting openness' of a jurisdiction (which is a function of commuting costs to other jurisdictions) matters as well.[1] If commuting costs are prohibitively high, residents work where they live and the standard model applies. However, if commuting costs are sufficiently low, workers are less likely to live where they work. In this set-up, changes in local public goods do not affect all workers in a jurisdiction but only the ones that also live in this jurisdiction. Whether this means, however, that employment is less sensitive to public good provision in commuting open jurisdictions depends on the importance of economies of scale. If wages change little, employment is less sensitive to public good provision in commuting open jurisdiction. However, if wages strongly increase with the number of workers, employment is more sensitive to public good provision in commuting open jurisdictions as local employment elasticities are larger in these jurisidcitions (Monte *et al.*, 2018).

Our paper assesses the empirical evidence for the impact of public good provision on local employment in commuting open jurisdictions. For identification, we focus on the closure of local grammar schools in East Germany after 2000. We believe this setting to be well suited for our research question: First, grammar schools are highly valued by German households as pupils finishing a grammar school in Germany obtain a university entrance qualification. In 2014, 41% of all pupils in East Germany went to a grammar school.[2] Moreover, the closure of a local grammar school is a very distinct and salient change in local public good provision. Second, the reason for the school closures was a general decline in the population in rural areas in East Germany after the German re-unification, and the decision on which grammar schools to close was made by the (hosting) state as it bears the main financial burden. This hedges against concerns that under-performing schools were closed or that the school closures had a substantial impact on local expenditures. In addition, as the school closures did not all happen

[1] The degree of commuting openness of a jurisdiction – i.e., the share of residents that out-commute and the share of workers that in-commute – is an equilibrium outcome. While commuting costs to other jurisdictions are an important aspect, equally important are household mobility/preferences and wage, public good and housing costs differentials.

[2] https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/BroschuereSchulenBlick0110018169004.pdf?__blob=publicationFile

in the same year, our setting allows us not only to implement a standard difference-in-differences (DiD) strategy but also to identify the impact of the school closures on local employment using only treated jurisdictions, which relaxes the assumption of exogenous school closures. Lastly, commuting plays an important role in our sample of jurisdictions. In the average jurisdiction, more than half of the local residents out-commute, and more than half of the local employees in-commute.

Our paper starts by outlining the theoretical impact of public good provision (directed to individuals rather than firms) on residents and employees using a Rosen–Roback type of local labor market model with commuting (see, for example, Monte *et al.* (2018)). While wages and house prices are endogenously determined, we assume public good provision to be independent of jurisdiction characteristics. This is implausible in the US context but largely true for the European and in particular the German contexts, where schools are almost entirely funded by the states. The theoretical framework predicts the two facts outlined above: First, the impact of local public good provision on local employment increases with the housing supply elasticity as more houses allow more workers to reside in a jurisdiction. Second, if economies of scale are unimportant (or wages decrease in the number of workers), the impact of local public good provision on local employment is predicted to decreases with the share of in-commuters since they do not value the public good provided in their workplace jurisdiction. However, if wages increase strongly with the number of workers, employment is more sensitive to public good provision in commuting open jurisdictions as employment elasticity are larger in those jurisdictions. In addition, the model highlights that with wages being a function of the number of workers in a jurisdiction, the impact of public good provision on residents and house prices depends on jurisdiction's commuting openness as well.

We test the theoretical predictions by quantifying the impact of local grammar school closures in East Germany after 2000 on the number of (employed) residents, local employees, house prices and land use using municipality level data (residents, employees, land use) for 1997 to 2008 and property level data for 2004 to 2008. Methodologically, we rely on an event study design using two different samples. Our first sample includes only jurisdictions in which the only grammar school was closed between 1992 and 2012. This set-up has the advantage that it does not require the assumption of exogenous school closures but only that the timing of the school closures was as good as random. It comes, however, at the cost of a rather small sample. Moreover, since all jurisdictions in the sample are treated, the estimates will only be unbiased if the treatment effects are sufficiently homogeneous (Abraham and Sun, 2018). Our second sample, which includes treated and control group jurisdictions, addresses this concern.

The downside of using the second sample is, however, that we only obtain an unbiased estimate if jurisdictions have been quasi-randomly selected into the treatment group. To increase the likelihood of this, we re-weight the control jurisdictions using entropy balancing (Hainmueller, 2012).

The two different samples provide largely similar results. We find that the school closures triggered a long-term decline in the number of residents by 11% and in the number of local employees by 5%. The smaller impact on local employment is consistent with the empirical observations that in-commuters are unaffected by the school closures and thus that wages change little in response to the school closures. Further, we find that house prices decreased by about 9.2% and land used for (residential) buildings by 2.4%. This implies a housing supply elasticity of around 0.26. If we assume a similar mobility of households as estimated in Ahlfeldt *et al.* (2015), our results suggest that 40% of the school closure burden is borne by residents and 60% by land owners.

Our work contributes to at least three streams of literature: First, we contribute to the literature on public policies in urban economies (Glaeser, 2008) and more specifically to the literature on local labor markets (e.g., Moretti, 2011; Notowidigdo, 2020) which employ and extend the spatial equilibrium model of Rosen (1979) and Roback (1982). Public policies may increase local employment by shifting labor demand or by shifting labor supply. While there is a large literature on demand shifters,[3] the literature on the impact of supply shifters – e.g., educational spending – on local employment is scarce. Further, existing work has either focused on relatively broad geographic areas (Mofidi and Stone, 1990; Dalenberg and Partridge, 1995) or used cross-sectional (within counties) as well as time-series variation for identification (Gabe and Bell, 2004). Although this literature is complemented by work that focuses on public good provision and individuals' location choice (Albouy and Lue, 2015; Buettner and Janeba, 2016), the implications for local employment are not clear as residents do not necessarily work where they live. A first contribution of our paper is thus that we analyze the impact of public good provision directed to households on local employment on a very detailed geographic level: the municipality level. Further, by exploiting changes in the availability of a local grammar school over time, our estimation strategy is robust to time-invariant heterogeneity between jurisdictions.

In addition, our work adds to our understanding of the role of commuting for local labor markets, which has recently been highlighted by Monte *et al.* (2018). They develop a tractable model that incorporates commuting (as well as trade) into a local labor market framework. The model predicts and the empirical analysis confirms that

[3]See, for example, Redding *et al.* (2011); Duranton and Turner (2012); Moeller and Zierer (2018); Busso *et al.* (2013) and Suàrez Serrato and Zidar (2016).

local employment elasticities are larger in more commuting open jurisdictions. The reason is that additional workers do not have to move but can simply commute to the workplace jurisdictions. Thus, public policies that shift the demand for workers have a stronger impact on local employment in more commuting open jurisdictions.[4] We contribute to this literature by addressing the question of whether a jurisdiction's commuting openness also affects the impact of public policies that act as supply shifters.

Second, our work links to studies that analyze the provision of local public goods in the context of local public finances. The papers most relevant for us are from the urban economics literature and combine a monocentric city model with commuting and the provision (and financing) of a public good (De Bartolome and Ross, 2003; de Bartolome and Ross, 2004; Hanushek and Yilmaz, 2007; De Bartolome and Ross, 2007; Hanushek and Yilmaz, 2013). Our paper shows important similarities to those studies but also differs in two important dimensions. First, we do not model workers of different classes ("poor/unskilled" and "rich/skilled") as we are not interested in explaining income differentials but only the residence and workplace location choice. Our setting is thus the interaction between multiple (small) towns with no social heterogeneity. Second, as spelled out above, our paper treats public good provision as independent from local economic conditions, which is a reasonable assumption for the European and in particular the German context.

Finally, we contribute to the literature on local real estate markets and to the literature on the capitalization of public policies (in particular educational spending) into property prices. We add to the former literature by providing an estimate of the housing supply elasticity for non-urban jurisdictions in East Germany in response to a negative public good shock. Our estimate of 0.26 is comparable to the work by Caldera and Johansson (2013) since housing is a durable good and thus housing supply elasticities are expected to be smaller for negative shocks (see, for example, Green *et al.* (2005); Glaeser and Gyourko (2005) and Notowidigdo (2020)).[5] We contribute to the latter literature by estimating the impact of the school closures on house prices and by illustrating that the estimated impact on house prices only reflects the change in the value of public good provision under very strict assumptions.

The remainder of this article is structured as follows. In section 2 we give the reader

[4]An additional insight of their work is that public infrastructure is not necessarily only a demand side shifter – as suggested above – as it also affects labor supply.

[5]Since housing supply elasticities are determined by regulatory as well as physical constraints, they differ between as well as within countries. For the US, Saiz (2010) estimates a long run (population weighted) elasticity of 1.75 and metropolitan specific estimates from 0.6 to 5.45. Suàrez Serrato and Zidar (2016) report a long run elasticity of 0.8. For Germany, housing supply estimates are only available at the national level: Caldera and Johansson (2013) report, for example, a long-run elasticity of 0.4.

a short overview of the institutional background and present in section 3 our theoretical framework that guides our empirical analysis. The empirical strategy and the data are presented in section 4, and section 5 contains the results. Finally, section 6 concludes.

2 Institutional Setting

In this section, we give the reader a brief overview of the German school system and of the reasons for and the process of the school closures in East Germany that we use for identification in the empirical part of this paper.

In Germany, the states are responsible for the provision of education (including the funding), but differences between states are relatively small. Children start going to primary school at the age of six or seven, and in most states they attend primary school for four years.[6] After that, pupils go either to a grammar school for eight or nine years (depending on the state) or to a secondary school for six years (*Realschule*).[7] Pupils finishing grammar school obtain a university entrance qualification, and pupils finishing secondary school a middle school degree.[8]

In general, the allocation of pupils to the two different school types is based on the primary school teacher recommendation, which depends largely on academic achievement in primary school. In three of the four states included in our analysis, the recommendation is binding. However, parents can and do petition, and then usually an entrance test decides. In the states included in our analysis, the share of pupils (in the relevant age range) going to a grammar school increased from round 33% in 2000 to 43% in 2008.[9]

For primary school choice, distance is important, but this is not the case for secondary/grammar school choice as children are allowed to apply to the school of their choice. Since in rural areas (e.g., in the jurisdictions analyzed in this paper), there is usually only one grammar school in close proximity and grammar schools differentiate themselves little, it is fair to say that most pupils in rural areas go to the closest

[6]In our sample of East German states, pupils only go to primary school in Brandenburg for six years – officially years five and six are called school-independent orientation level.

[7]Depending on the state, there are also comprehensive secondary schools (*Gesamtschule*), which offer children both types of school degrees. In our sample, these schools are only relevant in Brandenburg. However, despite this the share of pupils going to a grammar school in Brandenburg is very similar to the other states.

[8]Children can also finish secondary school after five years, in which case they receive a *Hauptschulabschluss*.

[9]See Table 21111-01-03-4-B on `www.regionalstatistik.de`, provided by the Federal Statistical Office. This increase was only possible because the overall number of pupils decreased.

grammar school/secondary school.[10]

Prior research (mainly for West Germany) has shown that pupils from households with a higher education level are more likely to go to a grammar school (Schulze *et al.* (2008)). While income seems to matters as well (but largely at the lower end of the distribution), it is less important than the education level of the parents (Schneider, 2004). Moreover, it should be noted that there is less income dispersion in East relative to West Germany[11], and our focus on the municipality level means also that there is little variation in income within a jurisdiction. We thus believe that the income of parents that wish to send their children to a grammar school differs – conditional on residence or workplace location – relatively little from the income of workers without children and the income of parents that do not wish to send their children to a grammar school.

States bear the main burden of the education system in Germany and are in particular responsible for funding all teachers, providing a per-pupil allowances and funding major investments. Hosting municipalities are responsible for school infrastructure and maintenance of school buildings, facilities and equipment. Hosting counties are responsible for pupil transportation. Neither counties nor municipalities, however, have to fund these costs on their own as the (state-specific) fiscal equalization schemes take the additional financing needs of hosting a school into account. This suggests that the quality of schools is, in contrast to the UK or US, largely independent of local economic conditions.

The school closures in East Germany after 2000 analyzed in this paper occurred due to an exceptional decrease in population during the 1990s and a resulting drop in the number of school-age children in the late 1990s and early 2000s. The reasons for the population decline are threefold. First, the reunification triggered a massive wave of migration from East to West.[12] As it was mostly the younger people moving, this would show a (delayed) direct effect on the number of pupils attending East German schools. Second, birthrates dropped in East Germany after the re-unification (likely due to economic uncertainty), which decreased the number of pupils thereafter (see, for example, Chevalier and Marie (2014)). Lastly, in the second half of the 1990s, there was a substantial migration from rural to urban areas with many young professionals

[10] In the event that demand exceeds supply in non-urban jurisdictions, classes are usually split if the difference is substantial.

[11] For example, in 2002, the share of households having a net income of more than 3,600 EURO per month was 25% in West Germany but only 10% in East Germany (Federal Statistical Office (2004): Wirtschaftsrechnung: Einnahmen und Ausgaben privater Haushalte, Fachserie 15 Reihe 1).

[12] Overall, about four million inhabitants (out of 16 million) moved from East to West Germany for at least some time.

and young families leaving.

The decision to close a particular school was entirely in the hands of the hosting state. While grammar school hosting municipalities certainly tried to actively lobby and exert political pressure, they did not have a final say in the process. In principle, we would expect the decision of which school to close to be based on the future demographic and economic outlook of the jurisdiction. However, based on our discussions with policy makers and studying the media coverage, we believe that many other considerations played a more prominent role. Among these were individual investment needs and status of a school, a particular historical status of a school and the availability of valid alternatives in reasonable proximity; political considerations were also of great importance. In addition, it is important to note that public administrators faced the same data limitations we face. Computer-based record keeping on all demographic information only started around 1997. Thus, the public decision makers in the late 90s and early 2000s did not have precise data to project future demographic developments on the municipality level. This suggests that the choice of which schools to close was related to expected employment growth only to a small degree, if at all.

3 Theoretical Motivation

In this section, we present a local labor model that will guide our empirical analysis. While the standard part of the model follows the work by Rosen (1979), Roback (1982) and Moretti (2011), commuting is introduced following the work by Monte *et al.* (2018). We are in a (static) world with N (atomic) households and (a large number of) jurisdictions, I. For simplicity, we normalize N to 1. There are three types of agents in the model: households, land owners and firms.

Households' Location Choice – Housing Demand and Labor Supply: Households choose their residence and workplace location and decide about housing and non-housing consumption. In the residence jurisdiction k, households face per-unit housing costs r_k and benefit from the local public good $v(g_k)$. In the workplace jurisdiction l, households provide inelastically one unit of labor and receive the wage w_l in return. Commuting costs for the residence-workplace location pair kl are given by κ_{kl}. Conditional on the residence and workplace location pair choice and prices, households decide on housing (H) and non-housing consumption (X). The price for non-housing consumption is the same in each city and is normalized to one.

Assuming a Cobb–Douglas utility function, household n chooses the residence and workplace location pair kl, consumption X_{kl} and housing H_{kl} to maximize[13]

$$U_{n,kl} = e_{n,kl}\ v(g_k) \left[\frac{X_{kl}}{(1-\gamma)}\right]^{(1-\gamma)} \left[\frac{H_{kl}}{\gamma}\right]^{\gamma} \text{ s.t. } \frac{w_l}{\kappa_{kl}} = X_{kl} + r_k H_{kl}$$

subject to the budget constraint, where γ is the share of household income spend on housing and $e_{n,kl}$ is household n's (idiosyncratic) preference for the location pair kl.[14]

Using the optimal quantities of housing and non-housing consumption, household n's indirect utility of choosing residence and workplace location pair kl is given by $V_{n,kl} = e_{n,kl}\ \frac{w_l\ v(g_k)}{\kappa_{kl}\ r_k^{\gamma}}$. To map the indirect utility into location pair choice probabilities, we assume that the $e_{n,kl}$'s are drawn from a Frećhet distribution with scale parameter $A_{kl}(= B_k C_l D_{kl})$ and shape parameter $\sigma > 0$. The scale parameter captures the average utility of living in k (B_k), of working in l (C_l) and of commuting between k and l (D_{kl}), and the shape parameter reflects households preference heterogeneity (or in other words the elasticity of households with respect to utility in monetary terms). If σ is large, household preferences are rather homogeneous and thus households are more responsive to wage, per-unit housing cost and public good provision changes; if σ is small, household preferences are more heterogeneous, and households are less responsive.

Based on this assumption, the probability that household n chooses the location pair kl is given by the indirect utility of choosing kl over the sum of the indirect utilities of any possible workplace-residence location pair (see equation (1)). Since we have normalized the number of households to 1, the probability that a particular household chooses location pair kl equals the number of households that choose to do so:

$$\pi_{kl} = N_{kl} = \frac{A_{kl}\left(\frac{w_l v(g_k)}{\kappa_{kl} r_k^{\gamma}}\right)^{\sigma}}{\sum_{i=1}^{I}\sum_{j=1}^{I} A_{ij}\left(\frac{w_j v(g_i)}{\kappa_{ij} r_i^{\gamma}}\right)^{\sigma}} \tag{1}$$

The number of households that live in jurisdictions k (e.g., the number of residents

[13] H and X have the subscript kl as both depend on wage income (which depends on the workplace jurisdiction) and housing costs (which depends on the residence jurisdiction).

[14] We include commuting costs in the household's budget constraint (and not the utility function, as for example do Monte *et al.* (2018)) as commuting involves real costs – e.g., for gasoline or public transport, which reduces the income that can be spend on consumption and housing. Moreover, commuting takes time, which may reduce working hours and therefore income.

in k), R_k, is then given by

$$R_k = \sum_{j=1}^{I} N_{kj} = \frac{\sum_{j=1}^{I} A_{kj}\left(\frac{w_j v(g_k)}{\kappa_{kj} r_k^{\gamma}}\right)^{\sigma}}{\sum_{i=1}^{I}\sum_{j=1}^{I} A_{ij}\left(\frac{w_j v(g_i)}{\kappa_{ij} r_i^{\gamma}}\right)^{\sigma}} \tag{2}$$

and the number of households that work in location k (e.g., the number of workers in k), W_k, is given by

$$W_k = \sum_{i=1}^{I} N_{ik} = \frac{\sum_{i=1}^{I} A_{ik}\left(\frac{w_k v(g_i)}{\kappa_{ik} r_c^{\gamma}}\right)^{\sigma}}{\sum_{i=1}^{I}\sum_{j=1}^{I} A_{ij}\left(\frac{w_j v(g_i)}{\kappa_{ij} r_i^{\gamma}}\right)^{\sigma}} \tag{3}$$

Given the focus of our paper, it is convenient to define $\lambda^R_{kl|k}$ as the share of residents in k that work in l ($\lambda^R_{kl|k} = \frac{N_{kl}}{R_k} = \frac{A_{kl}\kappa_{kl}^{-\sigma} w_l^{\sigma}}{\sum_{j=1}^{I} A_{kj}\kappa_{kj}^{-\sigma} w_j^{\sigma}}$) and $\lambda^W_{kl|l}$ as the share of workers in l that live in k ($\lambda^W_{kl|l} = \frac{N_{kl}}{W_k} = \frac{A_{kl}\kappa_{kl}^{-\sigma} r_k^{-\sigma\gamma} v(g_k)^{\sigma}}{\sum_{i=1}^{I} A_{il}\kappa_{il}^{-\sigma} r_i^{-\sigma\gamma} v(g_i)^{\sigma}}$).

Firms – Labor Demand: We assume that in each potential workplace jurisdiction l there is a representative firm that produces a traded output good Y_l for which we normalize the price to one. Firm profits are given by $\pi_l = Y_l - w_l W_l$ and the production function by $Y_l = F_l W_l^{(1+\alpha)}$, with F_l (> 0) as exogenously given jurisdiction-specific factor productivity, $1+\alpha$ as the output elasticity of labor and α as the elasticity of wages with respect to the number of workers.[15] Firm owners are assumed to be outside of the model. Firms maximize profits, which implies that the wage in jurisdiction l is given by

$$w_l = (1+\alpha) F_l W_l^{\alpha} \tag{4}$$

The wage in l increases (decreases) with the number of workers in l if the production function exhibits increasing (decreasing) returns to scale – e.g., $\alpha > 0$ ($\alpha < 0$).

Housing Market: Aggregated housing demand in jurisdiction k is the sum of housing demand of each potential household type (which follows from the utility maximization problem) multiplied by the number of households of this type in the jurisdiction:

$$HD_k = \left[\sum_{j=1}^{I} N_{kj}\frac{w_j}{\kappa_{kj}}\right]\gamma r_k^{-1} = R_k\left[\sum_{j=1}^{I} \lambda^R_{kj|k}\frac{w_j}{\kappa_{kj}}\right]\gamma r_k^{-1} \tag{5}$$

Housing demand in k increases with the amount of income spend on housing by

[15] We assume in the following that $\alpha\epsilon_w < 1$ where ϵ_w is the partial elasticity of households with respect to own wages (holding per-unit housing costs and public good provision fixed). This ensures that an initial public good shock cannot cause a wage increase such that all workers work in only one jurisdiction.

the residents in k and therefore with the number of residents (R_k), the average (net of commuting costs) wage of the residents ($\sum_{j=1}^{I} \lambda_{kj|k}^{R} \frac{w_j}{\kappa_{kj}}$) and the share of income spend on housing (γ). It decreases with the per-unit costs of housing (r_k). Since commuting decouples wages from housing costs in a particular jurisdiction, we define $\lambda_{kl|k}^{H}$ as the share of housing demand in k by households of type kl ($\lambda_{kl|k}^{H} = \frac{N_{kl|k} \frac{w_k}{\kappa_{kl}}}{\sum_{j=1}^{I} N_{kj|k} \frac{w_j}{\kappa_{kj}}}$).

Housing supply is assumed to be given by $HS_k = r_k^{\theta}$, where θ is the housing supply elasticity ($\theta \geq 0$). It increases with the per-unit cost of housing. Since in equilibrium housing demand equals housing supply, per-unit cost of housing in k is given by:

$$r_k = \left(R_k \gamma \sum_{j=1}^{I} \lambda_{kj|k}^{R} \frac{w_j}{\kappa_{kj}} \right)^{\frac{1}{1+\theta}} \tag{6}$$

Spatial Equilibrium: The spatial equilibrium is determined by labor market ($\sum_{j=1}^{I} W_j = 1$) and housing market clearing ($\sum_{i=1}^{I} R_i = 1$) and thus by equations (2), (3), (4) and (6). To highlight the mechanisms in the model, we start by illustrating the impact of changes in (ln) public good provision and changes in (ln) wages on (ln) number of households that choose location pair kl (e.g., with the total differential of (ln) N_{kl}, expressed as a function of public good provision and wages, for more details see Appendix B).

$$\text{d ln } N_{kl} = \epsilon_g \text{ d ln } v(g_k) + (\epsilon_w - \epsilon_r \lambda_{kl|k}^{H}) \text{ d ln } w_l - \epsilon_r \sum_{j \neq l}^{I} \lambda_{kj|k}^{H} \text{ d ln } w_j \tag{7}$$

The total differential shows that a marginal change in the valuation of public good provision in k causes a relative change in the number of households that choose kl by ϵ_g ($= \frac{\frac{1+\theta}{\gamma}}{\frac{1+\theta}{\gamma} + \sigma} \sigma$), the partial elasticity of households with respect to public good provision in the residence jurisdiction (taking into account resulting changes in house prices and the number of residents but holding wages in all jurisdictions fixed). ϵ_g increases with the incidence of public good provision on residents ($\frac{\frac{1+\theta}{\gamma}}{\frac{1+\theta}{\gamma} + \sigma}$) – which increases with the elasticity of housing supply with respect to income ($\frac{1+\theta}{\gamma}$) – and increases with the elasticity of households with respect to utility in monetary terms (σ).

In addition, a marginal change in (ln) wage in workplace jurisdiction l causes a change in the number of households of type kl by $\epsilon_w - \lambda_{kl|k}^{H} \epsilon_r$, the partial elasticity of households of type kl with respect to their "own" wages (holding public good provision and wages in all other jurisdictions constant). ϵ_w $\left(= \frac{\frac{1+\theta}{\gamma} + 1}{\frac{1+\theta}{\gamma} + \sigma} \sigma \right)$ is the partial elasticity of households with respect to wages in the workplace jurisdiction, holding average wages in the residence jurisdiction constant, and $-\epsilon_r (= \frac{\sigma}{\frac{1+\theta}{\gamma}+\sigma})$ the partial elasticity of households with respect to average wages of residents in the residence jurisdiction, holding the wage in the workplace jurisdiction constant. The elasticity of households of type kl with respect to their own wages is household type kl specific as these households only account for the share $\lambda^H_{kl}(< 1)$ of the housing demand in k. In other words, in commuting open jurisdiction, wage changes in the workplace jurisdiction capitalize less strongly into house prices in the residence jurisdictions as the wage changes are shared with all residents that do not work in this workplace jurisdiction. Therefore, in more commuting open jurisdictions, the elasticity of households with respect to their own wages is larger (see Monte *et al.*, 2018). Since households' indirect utility increases with the wage in the workplace jurisdiction and decrease with the cost of housing in the residence jurisdiction (which increases with wages), $\epsilon_w > 0$ and $-\epsilon_r < 0$. Since house prices cannot increase by more than wages, $\epsilon_w > \epsilon_r$. Thus, the number of households of type kl increases with the wage in the workplace jurisdiction l: $\epsilon_w - \lambda^H_{kl|k}\epsilon_r > 0$ and the effect is stronger in more commuting open jurisdictions (as $\lambda^H_{kl|k}$ is smaller in these jurisdictions). In addition and as before, the elasticity with respect to own wages increases with the elasticity of households with respect to utility in monetary terms (σ) and the elasticity of housing supply with respect to income ($\frac{1+\theta}{\gamma}$).

Lastly, a marginal change in (ln) wage in any other jurisdiction s ($\neq l$) changes the number of households that choose kl by -$\epsilon_r\lambda^H_{ks|k}$. The impact on average wages in the residence jurisdiction (which matters for the cost of housing) is given by the share of housing demand in k by households that work in l ($\lambda^H_{ks|k}$), and households respond to that by $-\epsilon_r$. Thus, while wage changes in the workplace jurisdiction capitalize less strongly into house prices in commuting open jurisdictions, wage changes in all non-workplace jurisdictions affect the residence-workplace location choice by affecting house prices in the residence jurisdictions. ϵ_r increases with the elasticity of households with respect to utility in monetary terms (σ) and decreases in the elasticity of housing supply with respect to income ($\frac{1+\theta}{\gamma}$).

Using the partial elasticities ϵ_g, ϵ_w, and ϵ_r, we derive the theoretical predictions for the impact of public good provision changes in k on the number of residents, number of workers, house prices and land use in jurisdiction k. To keep it tractable, we assume in the following that in all other jurisdictions $s(\neq k)$, the numbers of workers and residents, public good provision, wages and house prices are unchanged. To reflect this, we use partial derivative symbols. We illustrate at the end of the section how this assumption affects the predictions.

To derive the impact of a change in public good provision in k on the number of workers in k, we start from the fact that $W_k = \sum_{i=1}^{I} N_{ik}$ and therefore d ln W_k =

$\sum_{i=1}^{I} \lambda^W_{ik|k}$ d ln N_{ik}. Using equation (7) for the d ln N_{ik}'s and incorporating that wages are a function of the number of workers (d ln $w_k = \alpha$ d ln W_k) gives:

$$\frac{\partial \ln W_k}{\partial g_k} = \frac{\epsilon_g \lambda^W_{kk|k}}{1 - \alpha(\epsilon_w - \lambda^W_{kk|k}\epsilon_r \lambda^H_{kk|k})} \frac{\mathrm{d} \ln v(g_i)}{\partial g_k} \tag{8}$$

The numerator of the partial derivative captures the direct impact of public good provision changes in k on the workers in k. Since only workers that live in k are affected by public good provision, the impact is given by the share of workers in k that live in k ($\lambda^W_{kk|k}$) multiplied by the partial elasticity of households with respect to public good provision, ϵ_g, and the change in public good valuation ($\frac{\mathrm{d} \ln v(g_i)}{\partial g_k}$). The denominator captures the potential additional wage effect as the number of workers changes. This effect depends on commuting openness for two reasons. First, as only the wage for residents that work in k changes, the wage change capitalizes in the per-unit housing costs only by $\lambda^H_{kk|k}$ (see also above). Second, the change in per-unit cost of housing in k matters also only for workers that live in k ($\lambda^W_{kk|k}$). Since $\epsilon_w - \epsilon_r \lambda^H_{kk|k} > 0$, $\lambda^W_{kk|k} \leq 1$ and we assumed that $\alpha \epsilon_w < 1$, the number of workers increases with public good provision and the direct public good effect is exaggerated (mitigated) if wages increase (decrease) with the number of workers, $\alpha > 0$ (< 0). The impact of commuting openness on the partial derivative is ambiguous as commuting enters numerator and denominator and α can be positive or negative. If wages decrease in the number of workers ($\alpha < 0$) or increase only to a very small extent, the impact of public good provision will be smaller in more commuting open jurisdictions since in-commuters are only indirectly affected via wage changes. If, however, wages increase strongly in the number of workers ($\alpha > 0$), the larger employment elasticity in more commuting open jurisdictions (as wage changes capitalize less strongly in house prices) causes a stronger impact of public good provision on employment in more commuting open jurisdictions.

The partial derivative of the number of residents in k is (as $R_k = \sum_{j=1}^{I} N_{kj}$) given by

$$\frac{\partial \ln R_k}{\partial g_k} = \epsilon_g \frac{\mathrm{d} \ln v(g_i)}{\partial g_k} + \alpha \left[\lambda^R_{kk|k}\epsilon_w - \epsilon_r \lambda^H_{kk|k}\right] \frac{\partial \ln W_k}{\partial g_k} \tag{9}$$

and consists of three effects. The first captures the direct public good impact and is given by ϵ_g multiplied by the change in public good valuation, as all residents are directly affected. The second and third effects relate to the potential additional wage effect ($\alpha \frac{\partial \ln W_k}{\partial g_k}$). The second effect captures that the number of residents changes as the number of residents that work in k ($\lambda^R_{kk|k}$) changes by ϵ_w. This is the "own" wage effect on residents that work in k. The third effect relates to changes in the number of residents due to changes in the per-unit cost of housing in k. While all residents are affected by changes in per-unit housing costs, the effect size is only $\lambda^H_{kk|k}$ as the wage change affects housing price only via the share of housing demand of residents that work in k. Since the direct public good effect dominates the partial derivative (if wages decrease in the number of workers, otherwise but effects are positive), the number of residents increases with public good provision.[16] Importantly, commuting openness matters for the impact on residents only in the presence of wage effects, as otherwise the last two terms are zero.

The partial derivative for the per-unit cost of housing in k is given by

$$\frac{\partial \ln\, r_k}{\partial g_k} = \frac{1}{1+\theta}\left[\frac{\partial \ln\, R_k}{\partial g_k} + \lambda^H_{kk|k}\alpha\frac{\partial \ln\, W_k}{\partial g_k}\right] \tag{10}$$

and consists of two effects, the change in (ln) number of residents and the change in (ln) average wages due to wage changes ($\alpha\frac{\partial \ln\, W_k}{\partial g_k}$) for residents that work in k. If wages increase (decrease) with the number of workers, the wage effect increase (decreases) the direct public good effect. The sum of the the two effects, rescaled by the inverse of 1 plus housing supply elasticity, gives the partial derivates for the per-unit cost of housing. Thus, the larger the impact on residents and wages (assuming that wages increase in the number of workers), the stronger the impact on per-unit housing costs, and the larger the housing supply elasticity, the smaller the impact on per-unit housing costs.

Lastly, the partial derivative for housing supply is given by:

$$\frac{\partial \ln\, HS^S_k}{\partial g_k} = \theta\frac{\partial \ln\, r_k}{\partial g_k} \tag{11}$$

Summing up, the partial derivatives highlight that in the absence of wage effects, commuting (e.g., the difference between commuting open and commuting closed jurisdictions) only reduces the impact of public good provision on the number of workers (as only workers that live in the jurisdiction are affected) – the impact on residents and house prices (and therefore land use) is unchanged. In the presence of wage effects, also the impact on the number of residents and on house prices depends on jurisdiction's commuting openness. If wages decrease in the number of workers or increase in the number of workers but the difference in the elasticity of households with respect to their own wage in commuting open and commuting closed jurisdictions is not

[16]In commuting closed jurisdictions, the number of residents/workers increases with public good provision, even if wage decrease in the number of workers (see equation (8) with $\lambda^W_{kk|k} = 1$). Since the impact on workers (and therefore wages) is smaller in commuting open jurisdictions if wages decrease in the number of workers, the direct public good impact dominates the impact on the number of residents also in commuting open jurisdiction.

too large, the impact of public good provision on the number of workers will also be smaller in commuting open jurisdictions in the presence of wage effects; if the difference in the elasticity of households with respect to own wages is, however, sufficiently large, employment will respond stronger in commuting open jurisdictions.

Spatial Interactions: In the following, we discuss how potential spillovers into neighboring jurisdictions impact the partial derivatives derived. Since jurisdictions are linked via in- and out-commuting, the spatial interactions follow them. For simplicity, we focus on the direct spatial interactions – e.g., the effects that results from commuting from k to s (or vice versa). We start with out-commuting. Our model takes into account that out-commuters from k to s are directly affected by the changes in public good provision in k; it ignores, however, that – as out-commuters from k to s are part of the workforce in s – this could change wages in s (if $\alpha \neq 0$). Depending on the relationship between wages and workers, the wage change in s could exaggerate or mitigate the direct public good impact on out-commuters from k to s (similar to the wage effect on workers that live in k that is included in the model). In addition, as the wage change in s capitalizes into housing prices via workers in s that live in s, in-commuting from s to k may change as well. While the first effect may be relevant, the second effect is most likely less so as the direct public good effect on out-commuters is shared "twice" before going back to k: first with all workers in s that do not live in k and second with all residents in s that do not work in s.

Similar patterns exist for in-commuting. Our model incorporates that wage changes in k affect the number of in-commuters from s to k. It ignores, however, that the wage change would capitalize into housing prices in s, and thereby change the response by in-commuters to the wage change in k. The denominator of $\frac{\partial \ln W_k}{\partial g_k}$ would no longer be $\epsilon_w - \lambda^W_{kk|k}\epsilon_r\lambda^H_{kk|k}$ but $\epsilon_w - \epsilon_r \sum_{i=1} \lambda^W_{ik|i}\lambda^H_{ik|i}$. The change in house prices would in principle affect workers in s that live in s and therefore the number of workers and wages in s. This in turn would affect out-commuters from k to s. Since, however, similar to above, this second effect is shared twice (first with all residents in s that do not work in k and second with all workers in s that do not live in s), it is unlikely to be important. Moreover, it should be noted that if wages are not a function of the number of workers, the model outlined above captures all relevant parts.

4 Empirical Strategy, Data and Descriptive Statistics

We test the theoretical predictions outlined above using local grammar school closures in municipalities in the East German states of Brandenburg, Saxony-Anhalt, Mecklenburg-Vorpommern and Saxony between 1992 and 2012 using data from 1997 to 2008.[17] In particular, as we will outline below, we estimate the impact of the school closures on (ln) number of (employed) residents ($\frac{\partial R}{\partial g}$, equation (9)), on (ln) number of workers ($\frac{\partial W}{\partial g}$, equation (8)), on (ln) house prices ($\frac{\partial r}{\partial g}$, equation (10)) and on (ln) housing supply ($\frac{\partial HS}{\partial g}$, equation (11)).

Since wage data at the municipality level is not available to us, we infer the elasticity of wages with respect to the number of workers by estimating the impact of the school closures on (ln) number of in-commuters. Since in-commuters are only indirectly affected by the change in public good provision via wages changes (see equation (7)), the impact of the school closures on (ln) in-commuters is $\alpha \epsilon_w \frac{\partial W}{\partial g}$.

4.1 Municipality Level Estimations

Methodologically, we implement an event study design. Our estimation equation for the effects of the school closures on the number of residents and the number of local employees at the municipality level reads as follows:

$$\begin{aligned} Y_{i,s,t} &= \beta_{-6}\mathbb{1}[t \leq C_i - 6] + \sum_{s=-5}^{3} \beta_s \mathbb{1}[t = C_i + s] + \beta_4 \mathbb{1}[t \geq C_i + 4] \\ &\quad + \lambda_i + \nu_{t,s} + \gamma X_{i,t} + \epsilon_{i,t} \end{aligned} \tag{12}$$

$Y_{i,s,t}$ represents our dependent variable of interest in municipality i in state s at time t (for example (ln) residents or (ln) local employees). The data is provided by the Federal Employment Agency.

The first three terms on the right hand side capture the event study variables; all depend on the jurisdiction specific local grammar school closure year C_i. The first term

[17] Since several municipality mergers have happened in East Germany, we focus on jurisdictions that were not affected by mergers between 2003 and 2008. Results are largely similar when requiring that municipalities have not been affected at all by municipality mergers. We account for municipality mergers up to and including 2002 by treating jurisdictions that merged in later years as merged from the beginning of our sample period. To the extent that averages have to be calculated, we use the population in 1997 as weights. We exclude Thuringia from our sample as there is only one jurisdiction that had a grammar school in 1992, which was closed before 2012.

captures the long term effect before the school closure (6 and more years) and the third term the long term effect after the school closure (4 and more years). The term in the middle captures the year-specific effects in between (from -5 to +4 years). Our main variable of interest is the long-run impact of a grammar school closure on the different outcome variables, given by β_4. The underlying assumption here is that the materialization of the full impact of the school closure takes four years.[18] The remaining event study variables are included to allow us to inspect the timing of the impact. In particular, the point estimates for the indicator variables up to the school closure allow us to assess whether there is a common trend before the school closure. Since the school closures were announced up to two years before the actual school closure, according to interviews with municipality administrators and newspaper articles, we include six leads and do expect anticipation effects in the two years before the actual school closure. The remaining variables in our estimation equation give our control variables and the error term. The former include for all specifications at least municipality and state–year fixed effects. Since including municipality fixed effects means not all event study variables are identified, we set β_{-3} to zero (i.e., exclude it from the estimation). Thus all school closure effects are estimated relative to the effect three years before the school closure.

We estimate the event-study specification using two different samples. Our first sample includes only jurisdictions that had a grammar school in 1992 that was closed before 2012, since our estimation sample includes the years 1997 to 2008. We refer to this group as the treatment group. The data on the availability of local grammar schools is hand-collected and cross-checked with the number of grammar schools available at the county level for Germany. We focus on jurisdictions with one grammar school as we believe this provides the most credible variation for our analysis.[19] In total, we observe in our sample 25 towns that saw their only grammar school closed between 1992 and 2012, with the majority happening between 2002 and 2008. The geographical location of jurisdictions with one grammar school in 1992 that was closed before 2012 in our sample is depicted on the left hand side of Figure (1).

The main benefit of solely using treated jurisdictions to identify the impact of the school closures is that the validity of this strategy does not require the assumption of exogenous school closures but only that the timing of the school closures was as good as random. The downside is, however, that our final estimation sample in-

[18]Thus, the treatment effects are not dynamic as, for example, discussed in Borusyak and Jaravel (2017).

[19]In principle, school closures in cities with more than one grammar school could also be used if these closures cause excess demand for grammar schools. However, we do not have information on the number of students at the school level and thus are not able to account for school size (changes).

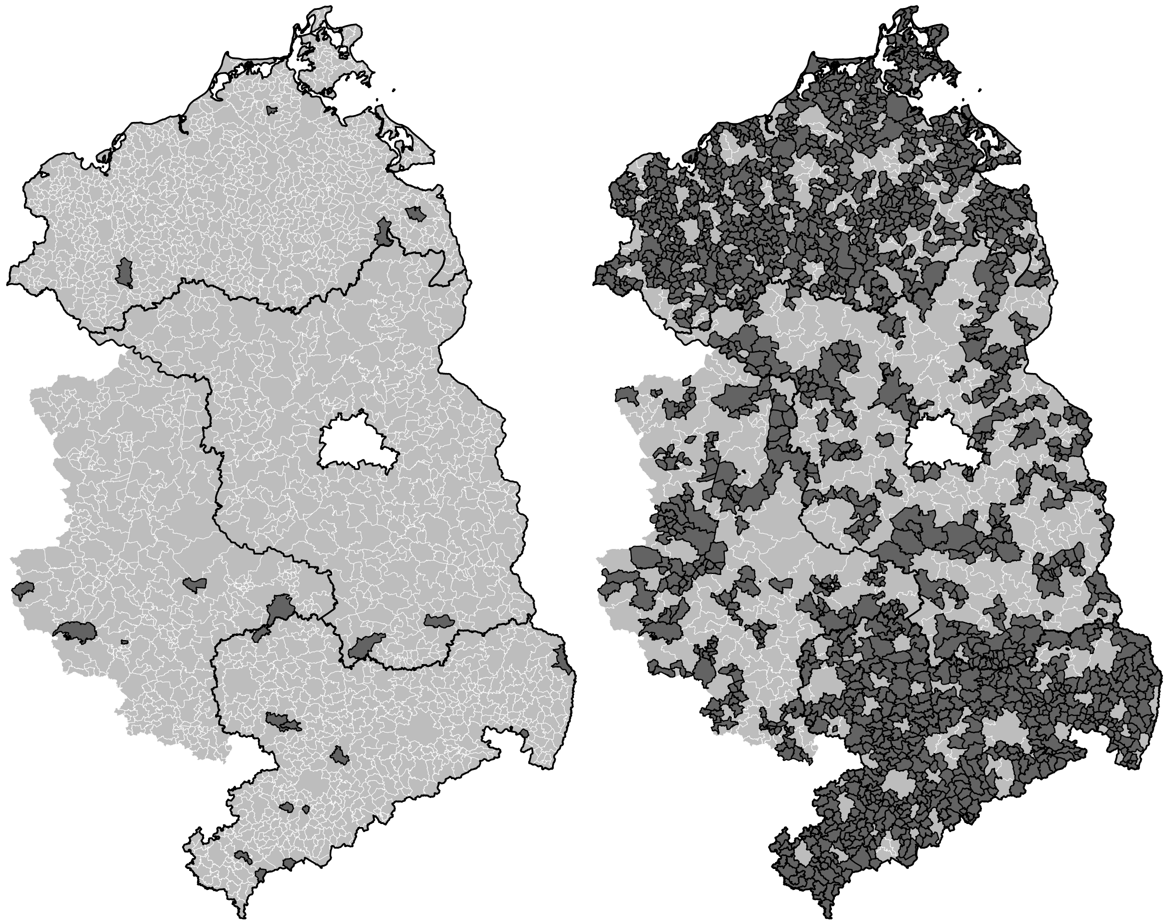
cludes only 25 unique jurisdictions and (due to missing information for some years) only 293 municipality–year observations. This small sample increases the risk of biased estimates in the event of heterogeneous treatment effects since all effects are solely identified based on the timing of the school closure (see Abraham and Sun (2018)). In particular, the before school closure effects of the jurisdictions with school closures toward the end of our estimation sample are solely identified relative to the post school closure effects of the jurisdictions with school closure at the beginning of our estimation period. In addition, the small sample limits the feasible set of control variables. In this specification, we thus control – in addition to the municipality and state–year fixed effects – only for (ln) population density in 1997 interacted with year dummies and (ln) local business tax and (ln) local property tax multiplier. We account for the two main tax instruments of municipalities to absorb any behavioral change of jurisdictions in response to the school closures. We report robust standard errors, clustered at the municipality level.

To address the small sample concern, our second estimation sample includes both treated and control group jurisdictions. Since our treatment group includes jurisdictions that had a grammar school in 1992 that was closed before 2012, the control group includes jurisdictions that had between 1992 and 2012 either always one grammar school or always none. This substantially larger sample size allows us to control for a wider range of control variables. Moreover, it ensures that the treatment effects are to a large extent identified relative to jurisdictions that are never treated. The downside of this approach is, however, that the DiD estimator for this sample will only provide unbiased estimates if there is a quasi-random selection of jurisdictions into the treatment group. As treated and control jurisdictions differ in size and commuting patterns, we re-weight our control group using a selection-on-observable framework to ensure that the characteristics of treatment and control group jurisdictions are similar at the beginning of our sample period. More precisely, we employ entropy balancing (Hainmueller (2012)). This approach obtains weights for each targeted moment of the balancing/matching variables for treatment and control group subject to the balancing constraints, and the resulting weights are then used in a weighted regression. The main advantage of this method compared to, for example, propensity matching is that the similarity of the distribution behind the mean of the matching variables is not just assumed but enforced in a constrained, nonlinear estimation approach.[20] To ensure a high degree of similarity between treatment and control groups, our set of balancing variables includes (ln) residents, (ln) in-commuters and (ln) out-commuters in 1998 and

[20]Moreover, entropy balancing does not disregard treatment and control group observations that cannot be matched exactly as does coarsened exact matching, which is another covariate balancing approach.

Figure 1: Spatial Distribution of Treatment and Control Group

(a) Treatment Group (b) Treatment and Control Group



Notes: The left hand side figure shows the location of jurisdictions in our sample that had a grammar school in 1992 that was closed before 2012. The right hand side figure shows the location of jurisdictions that had in 1992 either none or one grammar school . The white area in the middle is Berlin and is not included in the analysis. The darker lines show the borders of the states in East Germany.

the change in (ln) residents from 1998 to 1999; we match on the first two moments of their distribution.[21] Descriptive statistics for the balancing variables before and after entropy balancing are shown in Table (A1). As expected, treatment and control groups have identical means and standard deviations for the matching variables.

As the second approach samples 1,442 unique jurisdictions and 16,581 municipality–year observations, we include – in addition to the municipality and state-year fixed effects – the following control variables: (ln) population in 1997 and (ln) area size interacted with year dummies, (ln) distance to the next motorway interacted with

[21]Since the matching variables are missing for some jurisdictions, we impute missing values using county-average growth rates. The reader should note that we do not match in addition on (ln) local employees as this are given by (employed) residents less out-commuters plus in-commuters.

year dummies (to absorb commuting trends) as well as (ln) local business tax rate and (ln) local property tax multiplier. Since the weights from the entropy balancing are not known (but estimated), we report block-bootstrapped standard errors based on 500 replications.

To estimate the impact of the grammar school closures on land supply (on the municipality level), we rely on the following simplified specification:

$$Y_{i,s,t} = \beta \mathbb{1}[t \geq C_i] + \lambda_i + \nu_{t,s} + \gamma X_{i,t} + \epsilon_{i,t} \tag{13}$$

The reason is that the land use variables for Germany are only available for 1998, 2001, 2004 and 2007. Since this mean leads and lags cannot be estimated, we are only able to identify the (long run) effect of the school closures (β). The dependent variable in these specifications is either (ln) land used for buildings or the share of land used for buildings. While land used for buildings includes not only residential buildings (and related open space) but also commercial and public buildings, it has the benefit that it is available for 96% of our sample jurisdictions and that it exhibits a plausible variation over time.[22] Since we only have four years of data for the analysis and are not able to account for leads or lags, we only estimate the impact of the grammar school closures on land supply for our second sample, which includes both treated and matched control group observations. The set of control variables is the same as outlined above and, as before, we report block-bootstrapped standard errors.

Descriptive statistics for all variables at the municipality level are reported in Table (1). Overall, treatment and control group jurisdictions are suggested to be very similar after the entropy balancing.[23] In both groups, the number of employed residents is around 3,200 and the number of local employees around 2,650. Moreover, there are around 1,400 in-commuters and 1,950 out-commuters in the treatment and control group jurisdictions on average. The population is around 9,300, and there are around 320 children under the age of 6 in the average jurisdictions. Land used for buildings is around 300 hectares.

[22] We emphasize the latter here as land used for residential buildings is available for some jurisdictions in our sample. However, this variable exhibits implausible time trends in some states. For example, in Saxony land used for residential buildings grew by 400% between 1996 and 2008 despite a decline in the overall population (see Federal Statistical Office (2010)).

[23] The means and standard deviations still differ somewhat for the matching variables due to missing values and because we match on the logarithm of the matching variables.

Table 1: Descriptive Statistics for Local Employment Sample for 1998.

	Control Group				Treatment Group	
	Unbalanced		Entropy-Balanced			
	Mean	SD	Mean	SD	Mean	SD
Employed Residents	1,028	1,456	3,294	1,864	3,136	1,271
Local Employees	805	1,529	2,762	2,053	2,600	1,303
In-Commuters	493	908	1,471	1,036	1,381	752
Out-Commuters	714	867	1,990	998	1,917	784
Land Used for Buildings	118	124	292	155	323	191
Population Density	0.11	0.16	0.25	0.23	0.25	0.20
Population	2,872	4,098	9,257	5,486	9,473	3,937
Population 0-6 years	103	144	322	184	319	124
Local Business Tax Multiplier	312	47	337	41	347	36
Local Property Tax Multiplier	319	30	338	32	349	39
Observations	1,417		1,417		25	
(Max) Observations	16,288		16,288		293	

Notes: The table shows descriptive statistics (mean and standard deviation) for the variables at the municipality level for the treatment and control groups in 1998, before and after entropy balancing. The treatment group includes jurisdictions in which the only grammar school was closed between 1992 and 2012 and the control group jurisdictions that had between 1992 and 2012 either always no grammar school or always one. *Source:* Authors' calculations based on data from Federal Employment Agency and *Statistik Lokal*, 1997–2008.

4.2 Property-Level Estimations

To assess the impact of grammar school closures on house prices, we again employ an event study design. The estimation equation reads as follows:

$$\begin{aligned} Y_{k,i,s,t} &= \beta_{-4}\mathbb{1}[t \leq C_i - 4] + \sum_{s=-3}^{-1} \beta_s \mathbb{1}[t = C_i + s] + \beta_1 \mathbb{1}[t \geq C_i] \\ &\quad + \lambda_i + \nu_{t,s} + \gamma X_{k,i,t} + \epsilon_{i,t} \end{aligned} \tag{14}$$

The dependent variable is (ln) price of house k in municipality i in state s at time t. We estimate at the property level (and not as before at the municipality level) as this allows us to control for property characteristics more precisely. The house price data stems from Empirica AG, a real estate consulting company. It includes offer price data for residential properties from both newspapers and online ads, and is available to us for the years 2004 to 2008. Since the property markets in our sample of (non-urban) jurisdictions consist largely of houses, our baseline sample includes only purchase offers for single family homes, semi-detached houses and terraced houses. As houses are generally sold with land, we use (ln) price as the dependent variable and control for floor space/living space and amount of land (described in more detail below). In a robustness check, we also assess the impact of the school closures on flat prices.[24]

Since we only have five years of data and focus on purchase offers, which should

[24]While the database also covers rents, we focus on purchase prices as rents do not respond immediately to the full extent as prices do, and we only have five years of data.

respond immediately to the full extent[25], the event study design is captured by five variables. The first captures the long term effect before the school closure (4 and more years). The fifth variable captures the effect after the school closure, and the remaining three the effects related to 3, 2 and 1 year before the school closure, respectively. Since we again control for municipality and state-year fixed effects, we set β_{-3} to zero. Thus, all effects are identified relative to this year.

The house prices sample includes treated and control group observations to ensure that pre-trends are identified relative to jurisdictions that have not been treated. Since employing entropy balancing is rather difficult for this unbalanced sample at the property level, we include only jurisdictions into the control group that had a grammar school in 1992 that was not closed before 2012.[26] However, to account also in these specifications for a potential selection of jurisdictions into the treatment group, we include in all estimations all control variables on the municipality level and the balancing variables interacted with year dummies. In addition, we include the following property characteristics: (ln) amount of land (liner and squared), (ln) floor space (linear and squared) and the interaction between the two (linear and squared), type of property (single house, detached or terraced house), condition of the property (newly built, renovated, in need of renovation) and whether a basement, balcony and/or garage is available or not.[27] Lastly, we report robust standard errors that are clustered at the municipality level.

Descriptive statistics for this sample are shown in Table (2). The average property in the treatment group is located in a jurisdiction with 3,600 employed residents, which is somewhat larger than the municipality-level average. The same is also true for the control group jurisdictions. The average price of a property in the treatment group is 114,000 EURO and it comes with 135 square meter living space and 875 square meter open space. In the control group, houses are more expensive at 148,000 EURO, and while they have comparable figures for living space, they include less land. Moreover, there are less single-family homes in the control group jurisdictions but more semi-

[25] We assessed this in a robustness check and did not find any evidence for a delayed response. Results are available upon request.

[26] Results are very similar when using in addition jurisdictions that had no grammar school between 1992 and 2012. The common trend in these specifications is, however, somewhat less supported.

[27] Although we have additional information for some properties (e.g., location in the city, close to public transport or not, etc.), we decided not to use these as adverts differ strongly regarding the non-essential information included in the ad. Thus, we are concerned that additional property characteristics would not increase the precision of the estimates but only cause a selection bias. In the flat price estimation, we use the same set of property controls except for amount of land and the interaction between amount of land and floor space. We dropped all observations in the two samples in the top and bottom 1% of the distribution for price and floor space (as well as amount of land in the house price sample), to avoid outliers driving the results.

detached and terraced houses.

Table 2: Descriptive Statistics for House Price Sample for 2004 to 2008.

	Control Group		Treatment Group	
	Mean	SD	Mean	SD
Price	148,274	70,099	114,409	61,334
Amount of Land	7,68	698	875	771
Living Space	136	55	135	53
Single Family House	0.77	0.42	0.83	0.38
Detached House	0.13	0.34	0.10	0.30
Terraced House	0.09	0.29	0.07	0.25
Basement	0.21	0.41	0.20	0.40
Balcony	0.08	0.27	0.07	0.26
Garage	0.53	0.50	0.52	0.50
High Quality	0.12	0.32	0.08	0.28
Renovated	0.36	0.48	0.38	0.48
In Need of Renovation	0.20	0.40	0.22	0.41
Newly Build	0.07	0.25	0.06	0.23
(Employed) Residents (1998)	5,802	2,648	3,614	1,135
Local Employees (1998)	5,868	3,402	3,041	1,310
Out-Commuters (1998)	3,402	1,539	2,188	654
In-Commuters (1998)	3,466	2,088	1,611	760
Population Density (1997)	0.48	0.39	0.21	0.16
Local Business Tax Multiplier	349	41	356	32
Local Property Tax Multiplier	368	36	382	42
Observations	27,351		3,151	

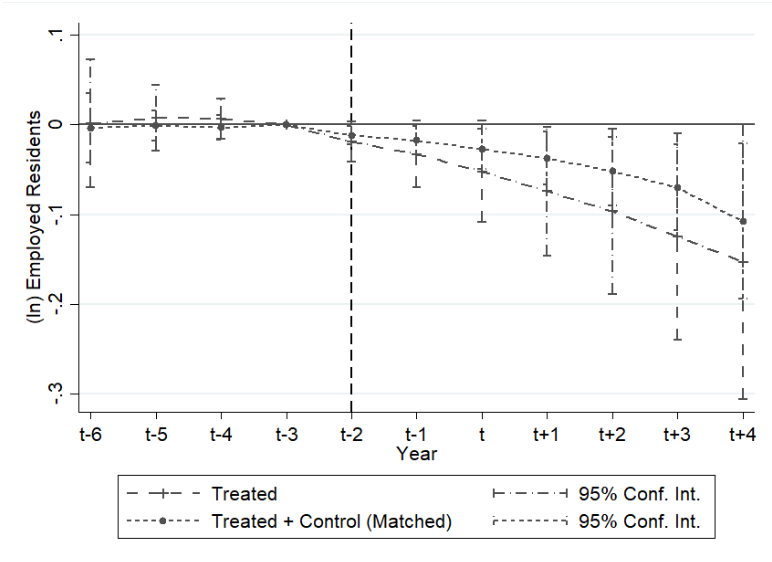
Notes: The table shows descriptive statistics (mean and standard deviation in parentheses) for the house price sample for the treatment and control groups. The treatment group includes jurisdictions in which the only grammar school was closed between 1992 and 2012 and the control group jurisdictions in which the only grammar school was not closed between 1992 and 2012. *Source:* Authors' calculations based on data from *Statistik Lokal* and Empirica AG, 2004–2008.

5 Results

In the following, we present the results. We start with the impact of the local grammar school closures on the number of (employed) residents and local employment before turning to the estimates for land use and house prices.

Residents: The point estimates for the 5 leads and 4 lags of the school closure variable based on the two different samples and using (ln) employed residents as dependent variable are shown in Figure (2). The horizontal axis represents the year relative to the year of the school closure (t), and the dashed line indicates the date when the school closure was announced, which is 2 years before the actual school closure. Up to the year of the (earliest) announcement, there is no evidence for a differential trend of treated jurisdictions (e.g., jurisdictions in which the school was closed in t), either when using only treated or when using treated and matched control group jurisdictions. Moreover, in both estimation samples, the number of (employed) residents decreases after the school closure announcement. The main difference between the two sub-samples is that the effects are more precisely estimated and somewhat smaller in absolute terms when using treated and matched control group observations, most likely as the sample size is larger. Quantitatively, the effect size is substantial as our results suggests a long term decrease in the number of (employed) residents by 11 to 15% (see col. (1) in Table (3)). To assess how important the entropy balancing is for the results using the large sample, col. (1) in Table (A2) in the Appendix shows the results when using treated and unmatched control group jurisdictions. The point estimate changes little, which is in line with our prior view that the decision on which schools to close was largely independent of the economic conditions in the grammar school hosting municipalities.

Figure 2: Estimation Results for (ln) Employed Residents



Notes: The figure depicts point estimates and 95% confidence intervals of the school closure event study variable using only treated jurisdictions and using treated and matched control group jurisdictions. The dependent variable is (ln) employed residents. *Source:* Authors' calculations based on data from Federal Employment Agency and *Statistik Lokal* 1997–2008.

To provide some more evidence that it is indeed the school closures that are driving our results, Figure (A1) in the Appendix shows that in particular children that are at an age of starting grammar school soon (6 to 10 years) leave the jurisdictions. Their number decreases by around 6 percentage points more compared to employed residents (see col. (5) in Table Table (3)).[28]

Given that our theoretical framework predicts that the impact on the number of residents depends on jurisdiction's commuting openness, if wages are a function of the number of residents, we split the sample using the share of workers that live in the

[28]We only use the second sample here as the age structure of the population is only available for 2000 to 2006 and 2008.

jurisdiction.[29] Table (A2) in the Appendix shows the results.[30] Col. (5) shows the long-run impact of the school closure in jurisdictions with a small share of workers that live in the jurisdiction (which should experience a smaller wage shock), and col. (6) shows the results for treated jurisdictions with a larger than median share of workers that live in the jurisdiction. For the latter group, the estimate is basically unchanged from our baseline estimate as shown in col. (1). For the former group, the effect is -9% – smaller in absolute terms than the baseline estimate. It is, however, also less precisely estimated. Thus, there is only weak evidence, if any, that wages increase with the number of workers and therefore that (ln) employed residents change in relative terms more strongly in jurisdictions with a larger share of workers that live in the jurisdiction.

Local Employees and Wages: We turn next to the results for (ln) employees, which are shown – when using only treated jurisdictions – in Figure (3). As for (ln) employed residents, there is no evidence for a differential trend in jurisdictions treated in t. After the school closure announcement, (ln) number of local employees declines. The long term effect is with -6% smaller in absolute terms than the impact on (ln) number of (employed) residents (see col. (2) in Table (3)). Given that the wage effect is suggested to be small (if it exists), the smaller relative impact on workers is in line with the prediction of the theoretical model that worker-residents are affected but not in-commuters. However, the estimates are not significantly different from zero, not even in the sample with treated and matched control group observations (see Figure (A2) and col. (2) in Table (3)).

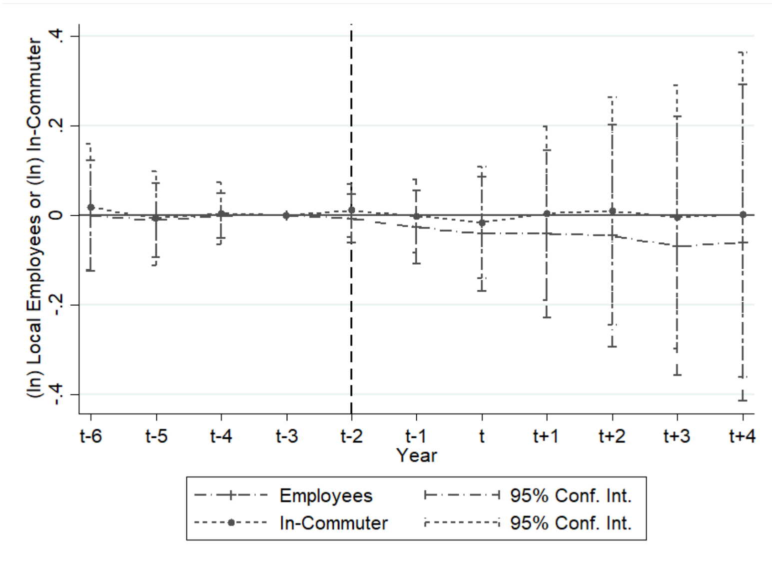
To assess whether wages are affected by the school closures, we inspect the impact of the school closures on (ln) in-commuters as they would only respond if wages change. The yearly point estimates – using only treated jurisdictions – are shown in Figure (3). There is little evidence for a change in the number of in-commuters. Our estimated long-run impact is literally zero (see col. (3) in Table (3)), which suggests that wages are independent of the number of workers, at least within our sample period. In the sample with treated and matched control group jurisdictions a negative long term effect is indicated (see Figure (A2) in the Appendix). However, the point estimate is not statistically different from zero and – when using the unmatched sample – again literally zero (see col. (3) in Table (A2)).

An alternative test for the impact of the school closures on wages is to compare the impact of the school closures on residents with the impact on out-commuters as the

[29]We use this criterion as it captures the strength of the wage effect and as we do not have information on the housing demand share of residents that work in the jurisdiction.

[30]Since the entropy balancing no longer converges, we rely on the unmatched sample for this analysis.

Figure 3: Estimation Results for (ln) Local Employees and (ln) In-Commuters



Notes: The figure depicts point estimates and 95% confidence intervals of the school closure event study variables using only treated jurisdictions. The dependent variable is (ln) local employees or (ln) in-commuters. *Source:* Authors' calculations based on data from Federal Employment Agency and *Statistik Lokal* 1997–2008.

latter should be unaffected by changes in wages in the jurisdiction (see Figure (A3) in the Appendix and col. (3) in Table (3).). Similar to before, there is no differential trend for the treated jurisdictions before the school closure announcement. Moreover, the number of out-commuters declines very similarly to the number of employed residents.

Summing up, we find a negative impact of the school closures on the number of residents as well as on local employees, the number of residents decreases, however, more strongly in response to the school closures than the number of local employees as in-commuters are not affected (as wages seem to respond little). This means that in commuting open jurisdictions part of the negative shock on local employment is exported to neighboring jurisdictions via out-commuters. To provided some direct evidence on this, we investigate how the closure of schools in neighboring jurisdictions affects the number of (employed) residents as well as the number of local employees using the same empirical strategy as for our main approach. The treatment group for this analysis includes jurisdictions that are within 10 km distance of a jurisdiction that had to close its only grammar school between 1992 and 2012. The control group includes jurisdictions for which the number of jurisdictions within 10 km distance of at least one grammar school was unchanged between 1992 and 2012.[31] The point estimates for

[31] We focus on jurisdictions that did not have a grammar school between 1992 and 2012 as this ensures that treatment and control group jurisdictions are comparable; in addition, we employ entropy balancing using the same set of matching variables as before.

Table 3: Estimation Results for (ln) Residents, Local Employees, In- and Out-Commuters

Dependent Variable (ln)	Employed Residents	Local Employees	In-Commuters	Out-Commuters	Residents 6–10 years/ Employed Residents
	(1)	(2)	(3)	(4)	(5)
Sample			Only Treatment		
School Closure (t+4, t+5,...)	-0.153*	-0.061	0.001	-0.132	
	(0.078)	(0.180)	(0.185)	(0.105)	
Observations	293	293	290	290	
Municipality FE	✓	✓	✓	✓	
State-Year FE	✓	✓	✓	✓	
(ln) Population Density-Year FE	✓	✓	✓	✓	

Sample			Treatment and Control (Matched)		
School Closure (t+4, t+5,...)	-0.108**	-0.053	-0.025	-0.096***	-0.061
	(0.044)	(0.079)	(0.079)	(0.033)	(0.056)
Observations	16,581	16,581	16,403	16,420	12,124
Municipality FE	✓	✓	✓	✓	✓
State-Year FE	✓	✓	✓	✓	✓
Control Variables	✓	✓	✓	✓	✓

Notes: The table shows point estimates and 95% confidence intervals for the long run (4 years and more) impact of grammar school closures on (ln) employed residents (col. (1)), (ln) local employees (col. (2)), (ln) in-commuters (col. (3)), (ln) out-commuters (col. (4)) and (ln) ratio of population aged between 6 and 10 years over employed residents using only treated jurisdictions (upper panel) and using treated and matched control group jurisdictions (with always zero or one grammar school between 1992 and 2012). Robust standard errors, clustered at the municipality level, (upper panel) or block-bootstrapped standard errors (lower panel) in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. *Source:* Authors' calculations based on data from Federal Employment Agency and *Statistik Lokal* 1997–2008.

the neighboring jurisdictions' school closure variables using (ln) number of residents and (ln) local employees as dependent variables are reported in Figure (A4) in the Appendix. The closure of the only grammar school in a neighboring jurisdiction decreases the number of local employees significantly, while it leaves the number of (employed) residents largely unchanged. In other words, commuters distributes the public good shock across space, which means that local shocks have a larger impact on the number of residents than on the number of local employees, but neighboring jurisdiction shocks have a larger impact on the number of local employees than on the number of residents.

House Prices: The results for the impact of the school closures on house prices are reported in Table (4). All specifications include our full set of municipality level control variables as well as the property controls.[32] Cols. (1) to (3) show the results for the impact of the school closure on (ln) house prices. While the impact in col. (1) is estimated relative to three years before the actual school closures, the impact in cols. (2) and (3) is estimated relative to three or more years before the actual closures as the effect for four or more years before is literally zero. Since the effect for t-2 is also close to zero and insignificant, no differential pre-trend is indicated. The estimate in cols. (1) and (2) are very similar and suggest a reduction in house prices of around 9% due to the school closures. In col. (3), we exclude the matching variables, which has, however, little impact on the point estimate. Col. (4) shows the results for (ln) flat prices. The point estimate is very similar but less precisely estimated, potentially due to the lower number of observations. Lastly, in cols. (5) and (6) we assess whether there is evidence that house prices respond more strongly in jurisdictions with a larger share of workers that live in this jurisdiction, which should be the case if wages are impacted by the school closures. Consistent with the results for in-commuters, we find that house prices changed very similarly in jurisdictions with a large and a small share of workers that live in the jurisdictions.

Land Use: The results for the impact of the school closures on land used for buildings (including open space) are shown in Table (5). In cols. (1) and (3), we use (ln) land used for buildings and in col. (2) the share of land used for buildings as the dependent variables. Cols. (1) and (2) show the results when using treated and matched control jurisdictions and col. (3) when using treated and unmatched control jurisdiction. In all specifications, our full set of control variables is included. The sample size is much smaller than for the other municipality level estimations as land use information

[32]The only difference to the other estimations is that we do not include (ln) population 1997 and (ln) area interacted with year dummies but instead include (ln) population density 1997 interacted with year dummies and control in addition for (ln) distance to Berlin interacted with year dummies since we control in these additional specifications only for the matching variables interacted with year dummies.

Table 4: Estimation Results for House Prices

Dependent Variable			(ln) Price of			
Sample		Houses		Flats	Houses	
					$\lambda^W_{kk\|k} > P50$	$\lambda^W_{kk\|k} < P50$
	(1)	(2)	(3)	(4)	(5)	(6)
School Closure (t-4,5..)	0.011			-0.007		
	(0.032)			(0.052)		
School Closure (t-2)	0.036	0.032	0.032	-0.013	0.059	0.008
	(0.033)	(0.027)	(0.026)	(0.047)	(0.039)	(0.037)
School Closure (t-1)	-0.035	-0.039	-0.035	-0.024	-0.105**	-0.009
	(0.056)	(0.055)	(0.056)	(0.054)	(0.052)	(0.077)
School Closure (t, t+1,...)	-0.088*	-0.092**	-0.087*	-0.088	-0.084	-0.099*
	(0.047)	(0.045)	(0.044)	(0.070)	(0.067)	(0.059)
Observations	30502	30502	30502	16693	11033	19469
Property Controls	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
State-Year FE	✓	✓	✓	✓	✓	✓
Control Variables	✓	✓	✓	✓	✓	✓
Matching Variables -Year FE	✓	✓		✓		

Notes: The table shows point estimates for the impact of grammar school closures on (ln) house prices (col. (1) to (3), (5) and (6)) and (ln) flat prices (col. (3)). All specifications include property controls, municipality and state-year fixed effects as well as (ln) population density, (ln) distance to next motorway and (ln) distance to Berlin interacted with year dummies and (ln) local business tax and (ln) local property tax multiplier. In addition, cols. (1), (2) and (4) include the matching variables interacted with year dummies. The sample for all specifications includes jurisdictions that had a grammar school in 1992 and that was closed (treatment) or not closed before 2012 (control group). Robust standard errors, clustered at the municipality level, in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. *Source:* Authors' calculations based on data from Federal Employment Agency, *Statistik Lokal* and Empirica AG 2004–2008.

is only available every three years. The effect size is very similar for the three different specifications and suggests a decrease in land used for buildings by around 2.4%. The estimate is, however, only precisely estimated when using the unmatched sample.[33]

Discussion of the Results and Implications: Our results provide strong evidence for a negative impact of school closures on population, house prices and land use. While we also find a negative impact on the number of workers, this is less precisely estimated. In addition, our results suggest, if at all, only small wage effects: in-commuters respond very little to the school closures and there is a similar impact on population and house prices in jurisdictions with both small and large shares of workers that live in the jurisdictions. Furthermore, we estimate a housing supply elasticity of 0.26 (ratio of change in land used for buildings of 2.4%, col. (2) in Table (5) to change in house prices of -9.2%, col. (2) in Table (4), see equation (11)). Based on equation (10) (which relates the impact on house prices to impact on population and wage change), this suggests that the wage effect accounts for less than 10% of the change in house prices, if it exists at all (using a long-run impact on population of 10.8 (see col. (1) in Table

[33]Due to the small sample, it is not possible to assess whether housing supply is different in jurisdictions with large vs. small shares of workers that live in the jurisdictions.

Table 5: Estimation Results for Land Use

Dependent Variable	(ln) Land Used for Buildings	Share of Land Used for Buildings in %	(ln) Land Used for Buildings
Sample	Control (Matched)	Treatment and ... Control (Matched)	Control (Unmatched)
	(1)	(2)	(3)
School Closure (t, t+1,...)	-0.024	-0.154	-0.025**
	(0.020)	(0.099)	(0.011)
Observations	5,581	5,581	5,581
Municipality FE	✓	✓	✓
State-Year FE	✓	✓	✓
Control Variables	✓	✓	✓
Implied Elasticity		0.022	

Notes: The table shows point estimates for the impact of grammar school closures on (ln) land used for buildings (cols. (1) and (3)) and share of land used for buildings (col. (2)). All specifications include municipality and state-year fixed effects as well as (ln) population in 1997, (ln) area and (ln) distance to the next highway interacted with year dummies and (ln) local business tax and (ln) local property tax multiplier. The sample for all three specifications includes jurisdictions that had a grammar school in 1992 that was closed before 2012 (treatment) and matched control group jurisdictions that had between 1992 and 2012 either always one grammar school or always none. (Block-)Bootstrapped standard errors (col. (1) and (2)) and robust standard errors, clustered at the municipality level (col. (3)), in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. *Source:* Authors' calculations based on data from *Statistik Lokal* 1998–2007.

3).[34]

We believe that there are at least three potential explanations for the small impact on wages, if it exists at all. First, there are many blue-collar workers in East Germany[35] and agglomeration gains mainly accrue to white-collar workers (Gould, 2007). Second, the considered jurisdictions are unlikely to form their own labor market. Thus, while the shock for the particular jurisdictions is large, the impact on the number of workers within the labor market is likely to be much smaller. Third, our empirical approach focuses on the short-run impact of local grammar school closure, and we can thus not rule out that in the long run, wages are a function of the number of residents.

While our estimated housing supply elasticity seems at the lower end, it compares quite favorably to the nationwide estimate of 0.4 for Germany by Caldera and Johansson (2013). In particular, as housing is a durable good, negative shocks are expected to result in a smaller change in the number of properties than positive shocks (Notowidigdo, 2020).

We now use our estimates to assess who bears the burden of the grammar school closures. Since our estimates do not allow us to infer the household mobility/household preference heterogeneity parameter σ (as the change in the valuation of public good provision due to the school closure is unknown), we illustrate the implied incidence for

[34] Re-arranging equation (10) shows that $\lambda^{H}_{kk|k}\alpha\frac{\ln W_k}{\partial g_k} = (1+\theta)\frac{\partial r_k}{\partial g_k} - \frac{\partial \ln R_k}{\partial g_k}$ and thus that the change in the average wage is $0.8\% = ((1+0.26)9.2\% - 10.8\%)$.

[35] Around 2000, 60% of the workforce in East Germany was low-skilled, compared to 20% in West Germany, see Figure 3 in Bilaine (2019).

different σ's. Our preferred estimate for σ is 6.7, from Ahlfeldt *et al.* (2015), as they focus on Germany and their model incorporates commuting.[36]

Assuming that wages do not respond at all means that (ln) number of residents changes by ϵ_g ($= \frac{\frac{1+\theta}{\gamma}}{\frac{1+\theta}{\gamma}+\sigma}\sigma$) times the change in (ln) public good valuation (see equation (9)). Using our estimated housing supply elasticity of 0.26 and assuming that households spend 30% of their income on housing and $\sigma = 6.7$ suggests that $\epsilon_r = 2.58$ and thus that the school closures reduced the value of public good provision by an equivalent of 4.2% of wages (=10.8%/2.58). We express the change in public good valuation in terms of wages as we set up the model such that a 1% change in the valuation of public good provision has the same impact on households' (indirect) utility as a 1% change in wages.[37]

If households were half as mobile (twice as mobile), the implied reduction in the value of public good provision in terms of wages would by 5.8% (3.4%). Since Ahlfeldt *et al.* (2015) focuses on 500 m grid data and the jurisdictions in our sample are larger than that, we believe that the change in public good valuation due to the school closures is around 5.8%.

In this case ($\sigma = 0.5*6.7$), the incidence on residents is 56% (change in public good valuation of -5.8% less implied change in rents of 0.3 * -8.7% (10.8%/1.26, based on the housing supply estimate of 0.26), relative to change in public good provision of -5.8%), and on land owners 44%.[38] This means house prices reflect less than half of the relative change in public good provision, due to housing supply adjustment in connection with household heterogeneity. As a comparison, if we assume that the opening of a grammar school increases public good valuation by 5.8%, and use a housing supply elasticity of

[36]Most of the work that reports estimates for the preference parameter is for the US. Suàrez Serrato and Zidar (2016) estimate a preference parameter of around 1.25 (1/0.8) using US state-level variation, Notowidigdo (2020) reports a preference parameter estimate of around 14 (1/0.07) and Diamond (2016) of 2 (college workers) and 4 (non-college workers) using metropolitan statistical area variation. None of the models incorporate commuting, however.

[37]Please note that in principle - when knowing the change in house prices - no estimate for the housing supply elasticity is required to uncover the change in public good valuation as from equation (1) follows that d ln $R_k = \sigma$d ln $v(g_k) - \sigma\gamma$d ln r_k (ignoring a potential general equilibrium effect). Thus, d ln $v(g_k) = \frac{1}{\sigma}$d ln $R_k + \gamma$d ln r_k, which gives -4.4% ($= \frac{-10.8\%}{6.7} - 0.3*9.2\%$). The small difference results from the fact that the observed house price change does not exactly equal the one implied by the estimated change in population and the estimated housing supply elasticity.

[38]If we assume a mobility of 6.7, the incidence share would be 39% (residents) and 61% (land owners). Please note that assuming a Cobb-Douglas utility function makes the housing demand more elastic. The difference is, however, relatively small in our case. The Cobb-Douglas housing demand elasticity is 4.2 ($\frac{1+\theta}{\gamma} = \frac{1.26}{0.3}$), while a more flexible specification ($\frac{1}{\gamma} + \theta$) would give a housing demand elasticity of 3.6.

2, this would suggest that residents bear 75% and landowners 25%.[39]

6 Conclusion

This article evaluates the role of public good provision directed to households on the spatial distribution of employment. Motivated by a local labor market model that includes commuting, we investigate the impact of grammar school closures on the number of residents, local employees, house prices and land use. We find a long run impact of the school closures on the number of residents of -11% and on the number of local employees of -5%. This differential impact is consistent with the theoretical prediction that in the absence of wage effects (which is supported by our data) in-commuters are unaffected by local public good provision. Our results thus point to the fact that commuters distribute public good shocks across space and thus that public good provision directed to households may be an ineffective policy tool for stimulating employment in commuting open jurisdictions.

[39] This is calculated as follows: Changing the housing supply elasticity to 2 means $\epsilon_g = 2.51$, population changes therefore by 14.5% (2.51 * 5.8% (change public good valuation)) and housing costs by 4.8% (14.5% * 1/(1+2)).

References

Abraham, S. and Sun, L. (2018). *Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects*. Papers 1804.05785, arXiv.org.

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M. and Wolf, N. (2015). The economics of density: Evidence from the berlin wall. *Econometrica*, **83** (6), 2127–2189.

Albouy, D. and Lue, B. (2015). Driving to opportunity: Local rents, wages, commuting, and sub-metropolitan quality of life. *Journal of Urban Economics*, **89**, 74 – 92.

Bilaine, K. (2019). *Low-skilled workers in Germany - low wages, precarious conditions*. Tech. rep., Bertelsmann Stiftung, Policy Brief 2019/05.

Borusyak, K. and Jaravel, X. (2017). Revisiting event study designs. *Mimeo*.

Buettner, T. and Janeba, E. (2016). City competition for the creative class. *Journal of Cultural Economics*, **40** (4), 413–451.

Busso, M., Gregory, J. and Kline, P. (2013). Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review*, **103** (2), 897–947.

Caldera, A. and Johansson, s. (2013). The price responsiveness of housing supply in OECD countries. *Journal of Housing Economics*, **22** (3), 231–249.

Chevalier, A. and Marie, O. (2014). Economic uncertainty, parental selection and the criminal activity of the "children of the wall". *CEP Discussion Paper No. 1256*.

Dalenberg, D. R. and Partridge, M. D. (1995). The effects of taxes, expenditures, and public infrastructure on metropolitan area employment. *Journal of Regional Science*, **35** (4), 617–640.

De Bartolome, C. A. and Ross, S. L. (2003). Equilibria with local governments and commuting: income sorting vs income mixing. *Journal of Urban Economics*, **54** (1), 1–20.

de Bartolome, C. A. and Ross, S. L. (2004). Who's in charge of the central city? the conflict between efficiency and equity in the design of a metropolitan area. *Journal of Urban Economics*, **56** (3), 458–483.

De Bartolome, C. A. and Ross, S. L. (2007). Community income distributions in a metropolitan area. *Journal of Urban Economics*, **61** (3), 496–518.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, **106** (3), 479–524.

Duranton, G. and Turner, M. (2012). Urban growth and transportation. *Review of Economic Studies*, **79** (4), 1407–1440.

Federal Statistical Office (2010). Land use statistics [Bodenflaeche nach Art der tatsaechlichen Nutzung]. *Fachserie 3 Reihe 5.1.*

Gabe, T. and Bell, K. P. (2004). Tradeoffs between local taxes and government spending as determinants of business location. *Journal of Regional Science*, **44** (1), 21–41.

Glaeser, E. L. (2008). *Cities, agglomeration, and spatial equilibrium.* Oxford University Press.

— and Gyourko, J. (2005). Urban decline and durable housing. *Journal of Political Economy*, **113** (2), 345–375.

Gould, E. (2007). Cities, workers, and wages: A structural analysis of the urban wage premium. *Review of Economic Studies*, **74** (2), 477–506.

Green, R. K., Malpezzi, S. and Mayo, S. K. (2005). Metropolitan-specific estimates of the price elasticity of supply of housing, and their sources. *American Economic Review*, **95** (2), 334–339.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, **20** (1), 25â€“46.

Hanushek, E. and Yilmaz, K. (2007). The complementarity of tiebout and alonso. *Journal of Housing Economics*, **16** (2), 243–261.

Hanushek, E. A. and Yilmaz, K. (2013). Schools and location: Tiebout, alonso, and governmental finance policy. *Journal of Public Economic Theory*, **15** (6), 829–855.

Moeller, J. and Zierer, M. (2018). Autobahns and jobs: A regional study using historical instrumental variables. *Journal of Urban Economics*, **103**, 18 – 33.

Mofidi, A. and Stone, J. (1990). Do state and local taxes affect economic growth? *The Review of Economics and Statistics*, **72** (4), 686–91.

Monte, F., Redding, S. J. and Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. *American Economic Review*, **108** (12), 3855–90.

Moretti, E. (2011). Local labor markets. *Handbook of Labor Economics*, **4**, 1237–1313.

Notowidigdo, M. J. (2020). The incidence of local labor demand shocks. *Journal of Labor Economics*, **38** (3), 000–000.

Redding, S. J., Sturm, D. M. and Wolf, N. (2011). History and industry location: Evidence from german airports. *Review of Economics and Statistics*, **93** (3), 814–831.

Roback, J. (1982). Wages, rents and the quality of life. *Journal of Political Economy*, **90** (6), 1257–1278.

Rosen, S. (1979). Wage-based indexes of urban quality of life. In P. Mieszkowsky and M. Straszheim (eds.), *Current issues in urban economics*, John Hopkins University Press.

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, **125** (3), 1253–1296.

Schneider, T. (2004). Der einfluss des einkommens der eltern auf die schulwahl [the influence of parental income on school choice]. *Zeitschrift fÃ¼r Soziologie*, **33** (6), 471–492.

Schulze, A., Unger, R. and Hradil, S. (2008). Bildungschancen und Lernbedigungen an Wiesbadener Grundschulen am Uebergang zur Sekundarstufe I. Projekt- und Ergebnisbericht zur Vollerhebung der GrundschuelerInnen der 4. Klasse im Schuljahr 2006/2007. *Projektgruppe Sozialbericht zur Bildungsbeteiligung. Amt fuer Soziale Arbeit, Abteilung Grundsatz und PLanung, Landeshauptstadt Wiesbaden.*

Suàrez Serrato, J. C. and Zidar, O. (2016). Who benefits from state corporate tax cuts? A local labor markets approach with heterogeneous firms. *American Economic Review*, **106** (9), 2582–2624.

Appendix A: Additional Tables and Figures

Table A1: Descriptive Statistics for Balancing Variables before and after Entropy Balancing

	Control Group				Treatment Group	
	Unbalanced		Entropy-Balanced			
	Mean	SD	Mean	SD	Mean	SD
(ln) Employed Residents (1998)	6.29	1.10	7.96	0.53	7.96	0.53
(ln) Out-Commuter (1998)	6.05	1.01	7.47	0.49	7.47	0.49
(ln) In-Commuter (1998)	5.15	1.49	7.04	0.71	7.04	0.71
d.(ln) Employed Residents (1998 to 1999)	0.01	0.07	-0.01	0.04	-0.01	0.04
Non-Balancing Variable						
(ln) Local Employees (1998)	5.63	1.47	7.68	0.69	7.71	0.71
Observations	1,417		1,417		25	

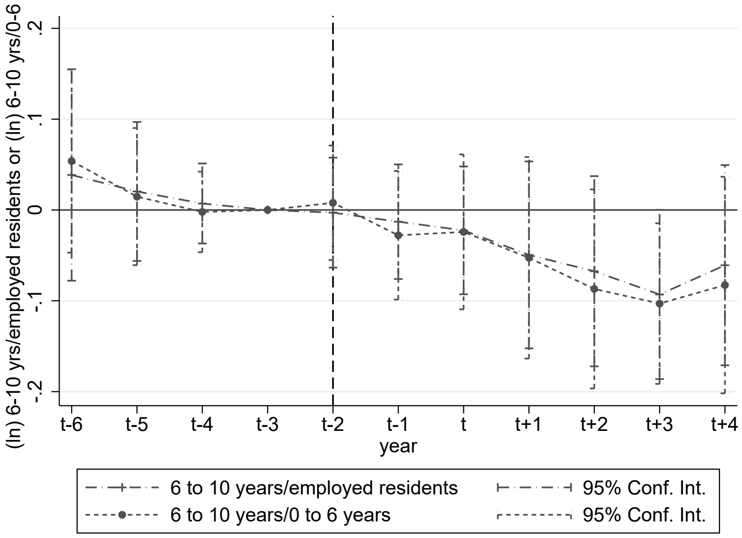
Notes: The table shows descriptive statistics (mean and standard deviation) for the balancing variables before and after entropy balancing. Missing values in the balancing variables are imputed using county-averages. The control group includes jurisdictions that had between 1992 and 2012 either always one grammar school or always none. The treatment group includes jurisdictions with a grammar school in 1992 that was closed until 2012. *Source:* Authors' calculations based on data from Federal Employment Agency and*Statistik Lokal* 1998 and 1999.

Table A2: Results for Sensitivity Analysis: Treated and Unmatched Control Jurisdictions

Dependent Variable (ln)	Employed Residents	Local Employees	In-Commuter	Out-Commuter	Employed Residents $\lambda^W_{kk\|k} < P50$	Employed Residents $\lambda^W_{kk\|k} > P50$
	(1)	(2)	(3)	(4)	(5)	(6)
School Closure (t+4, t+5,...)	-0.115***	-0.060	-0.007	-0.091***	-0.087*	-0.112***
	(0.043)	(0.069)	(0.066)	(0.027)	(0.048)	(0.038)
Observations	16592	16592	16412	16431	11596	4996
Municipality FE	✓	✓	✓	✓	✓	✓
State-Year FE	✓	✓	✓	✓	✓	✓
(ln) Population Density-Year FE	✓	✓	✓	✓	✓	✓

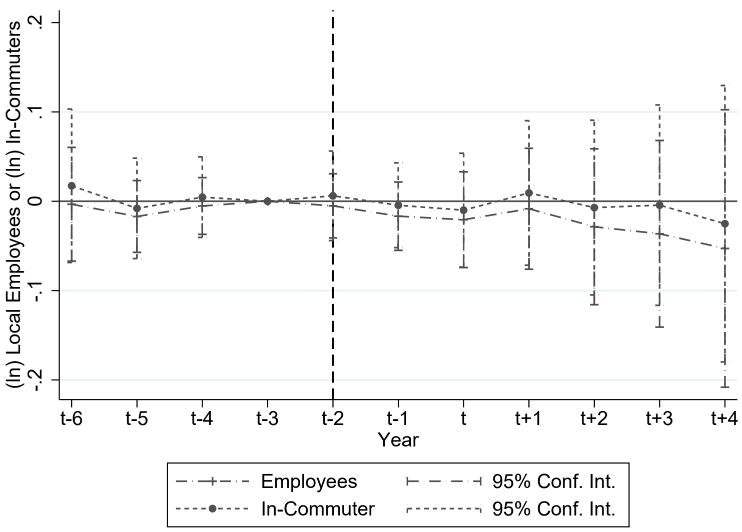
Notes: The table shows results of sensitivity analysis for the long run (4 years and more) impact of grammar school closures using treated and unmatched control group jurisdictions. The dependent variables are (ln) employed residents (cols. (1), (5) and (6)), (ln) local employees (col. (2)), (ln) in-commuters (col. (3)), (ln) out-commuters (col. (4)). Col. (1) to (4) use the full sample, the sample in col. (5) includes only jurisdictions with a small share of workers that live in the jurisdiction (smaller than median for treated jurisdictions) and col. (6) only jurisdictions with a large share of workers that live in the jurisdiction (larger than median for treated jurisdictions). Robust standard errors, clustered at the municipality level, in parenthesis. *, **, *** denote significance at the 10%, 5% and 1% level. *Source:* Authors' calculations based on data from Federal Employment Agency and *Statistik Lokal* 1997–2008.

Figure A1: Estimation Results for (ln) Population 6-10 years over Employed Residents and (ln) Population 6-10 years over 0-6 years: Treated and Matched Control Jurisdictions



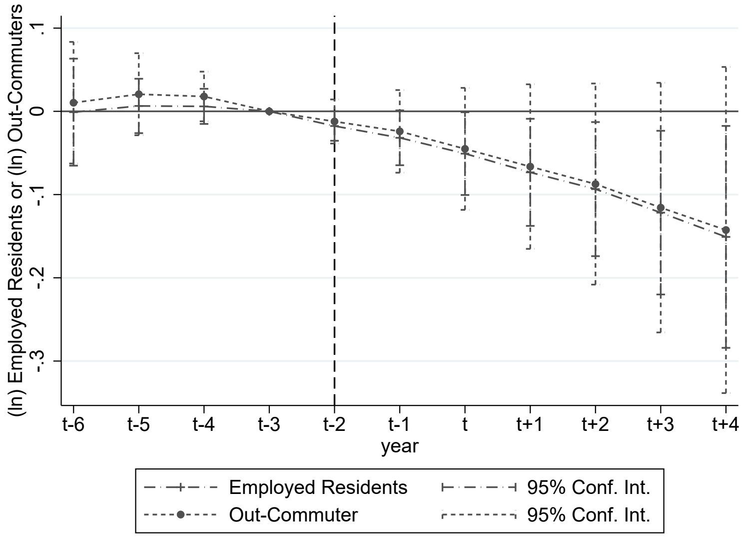
Notes: The figure depicts point estimates and 95% confidence intervals of the school closure variables using treated and matched control group jurisdictions. The dependent variable is (ln) population aged 6 to 10 years over employed residents or (ln) population aged 6 to 10 years over 0 to 6 years. *Source:* Authors' calculations based on data from Federal Employment Agency and *Statistik Lokal* 1997−2008.

Figure A2: Estimation Results for (ln) Workers and (ln) In-Commuters: Treated and Matched Control Jurisdictions



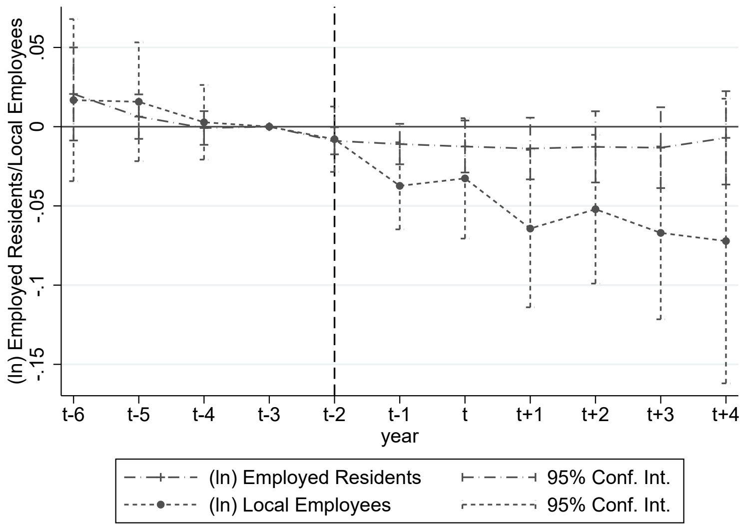
Notes: The figure depicts point estimates and 95% confidence intervals of the school closure event study variables using treated and matched control group jurisdictions. The dependent variable is (ln) local employees or (ln) in-commuters. *Source:* Authors' calculations based on data from*Statistik Lokal* 2000 − 2006 and 2008.

Figure A3: Estimation Results for (ln) Employed Residents and (ln) Out-Commuter



Notes: The figure depicts point estimates and 95% confidence intervals of the school closure event study variables using only treated jurisdictions. The dependent variable is (ln) employed residents or (ln) out-commuters. *Source:* Authors' calculations based on data from Federal Employment Agency and *Statistik Lokal* 1997 − 2008.

Figure A4: School Closures in Neighboring Jurisdictions: Estimation Results for (ln) Residents and Local Employees



Notes: The figure depicts point estimates and 95% confidence intervals of the neighboring jurisdiction school closure event study variable. The sample includes all jurisdictions that had no grammar school between 1992 and 2012 and for which the number of jurisdictions within 10km radius with at least one grammar school was unchanged (control) or reduced by one (treatment group). The control group jurisdictions are re-weighted using entropy-balancing. The dependent variable is (ln) employed residents or (ln) local employees. *Source:* Authors' calculations based on data from Federal Employment Agency and *Statistik Lokal* 1997 − 2008.

Appendix B: Derivation of Partial Derivatives for Changes in Public Good Provision

In this Appendix, we show step by step how we derived the partial derivatives. The spatial equilibrium in our theoretical model is given by labor and housing market clearing. To highlight the transmission channels at work, we start by focusing on the determinants of relative changes in the number of households that choose the location pair kl.

The number of households that choose kl is given by (see also equation (1)):

$$N_{kl} = A_{kl} \left(\frac{w_l v(g_k)}{\kappa_{kl} r_k^{\gamma}} \right)^{\sigma} U^{-1} \tag{B1}$$

with $U (= \sum_{i=1}^{I} \sum_{j=1}^{I} A_{ij} \left(\frac{w_j v(g_i)}{\kappa_{ij} r_i^{\gamma}} \right)^{\sigma})$ as the sum of the indirect utility for all possible residence-workplace location pairs.

Replacing the costs of per-unit housing, r_k, with equation (6) in the main paper gives:

$$N_{kl} = A_{kl} \left(\frac{w_l}{\kappa_{kl}} \right)^{\sigma} v(g_k)^{\sigma} \left[R_k \gamma \sum_{j=1}^{I} \lambda_{kj|k}^{R} \frac{w_j}{\kappa_{kj}} \right]^{-\frac{\gamma\sigma}{1+\theta}} U^{-1} \tag{B2}$$

Since $N_{kl|k}^{R} = \lambda_{kl|k}^{R} R_k$ and the denominators for λ^R (see page 9) are the same, this can be re-written as:

$$N_{kl}^{1+\frac{\gamma\sigma}{1+\theta}} = A_{kl}^{1-\frac{\gamma\sigma}{1+\theta}} \left(\frac{w_l}{\kappa_{kl}} \right)^{\sigma+\frac{\gamma\sigma}{1+\theta}} v(g_k)^{\sigma} \left[\sum_{j=1}^{I} A_{kj} \left(\frac{w_j}{\kappa_{kj}} \right)^{1+\sigma} \right]^{-\frac{\gamma\sigma}{1+\theta}} \gamma^{-\frac{\gamma\sigma}{1+\theta}} U^{-1} \tag{B3}$$

and then re-arranged to:

$$N_{kl} = A_{kl}^{\frac{1+\theta-\gamma\sigma}{1+\theta+\gamma\sigma}} \left(\frac{w_l}{\kappa_{kl}} \right)^{\epsilon_w} v(g_k)^{\epsilon_g} \left[\sum_{j=1}^{I} A_{kj} \left(\frac{w_j}{\kappa_{kj}} \right)^{1+\sigma} \right]^{-\epsilon_r} U^{-\frac{1+\theta}{1+\theta+\gamma\sigma}} \gamma^{-\frac{\gamma\sigma}{1+\theta+\gamma\sigma}} \tag{B4}$$

with $\epsilon_g = \frac{\frac{1+\theta}{\gamma}}{\frac{1+\theta}{\gamma}+\sigma}\sigma$ as the partial elasticity of households with respect to public good provision in the residence jurisdiction (holding wages in all jurisdictions constant), $\epsilon_w = \frac{\frac{1+\theta}{\gamma}+1}{\frac{1+\theta}{\gamma}+\sigma}\sigma$ as the partial elasticity of households with respect to own wage changes (holding public good provision and the average income of residence in the residence jurisdiction constant) and $-\epsilon_r = -\frac{\sigma}{\frac{1+\theta}{\gamma}+\sigma}$ as the partial elasticity of households with respect to changes in average wage of residents. ϵ_g (which equals $\epsilon_w - \epsilon_r$) is the incidence of utility changes in monetary terms on households, $\frac{\frac{1+\theta}{\gamma}}{\frac{1+\theta}{\gamma}+\sigma}$, multiplied with the elasticity of households with respect to utility in monetary terms σ.

Taking the total differential of ln N_{kl} gives then:

$$d \text{ ln } N_{kl} = \epsilon_w \text{d ln } w_l + \epsilon_g \text{d ln } v(g_k) - \epsilon_r \sum_{j=1}^{I} \lambda^H_{kj|k} \text{ d ln } w_j - \frac{1+\theta}{1+\theta+\gamma\sigma} d \text{ ln } U \quad \text{(B5)}$$

Since U and therefore d ln U is the same for all households types, we ignore this general equilibrium effect in the following as it cannot be identified empirically when including year fixed effects. Moreover, if the number of jurisdictions is large, changes in one jurisdictions have little impact on U. This gives then equation (7):

$$\text{d ln } N_{kl} = \epsilon_g \text{d ln } v(g_k) + (\epsilon_w - \lambda^H_{kl|k}) \text{d ln } w_l - \epsilon_r \sum_{j \neq l}^{I} \lambda^H_{kj|k} \text{ d ln } w_j \quad \text{(B6)}$$

Since $W_k = \sum_{i=1}^{I} N_{ik}$, the total differential of ln W_k is given by:

$$\text{d ln } W_k = \sum_{i=1}^{I} \lambda^W_{ik|k} \text{d ln} N_{ik} \quad \text{(B7)}$$

Assuming that d ln w_s = d ln $v(g_s)$ = d ln $r_s = 0$ for all $s \neq k$ gives then:

$$\partial \text{ ln } W_k = \lambda^W_{kk|k} \epsilon_g \partial \text{ ln } v(g_k) + (\epsilon_w - \lambda^W_{kk|k} \epsilon_r \lambda^H_{kk|k}) \partial \text{ln } w_k \quad \text{(B8)}$$

where we use ∂ to reflect that we ignore any changes in other jurisdictions. Since d ln $w_k = \alpha$ d ln W_k, the equation can be re-arranged to:

$$\partial \text{ ln } W_k = \frac{\lambda^W_{kk|k} \epsilon_g}{1 - \alpha(\epsilon_w - \lambda^W_{kk|k} \epsilon_r \lambda^H_{kk|k})} \partial \text{ ln } v(g_k) \quad \text{(B9)}$$

Dividing by ∂g_k gives then equation (8) in the paper. Proceeding in the same way for the number of residents in k gives:

$$\partial \text{ ln } R_k = \epsilon_g \partial \text{ ln } v(g_k) + \alpha(\lambda^R_{kk|k} \epsilon_w - \epsilon_r \lambda^H_{kk|k}) \partial \text{ ln } W_k \quad \text{(B10)}$$

Dividing by ∂g_k give then equation (9) in the paper. Lastly, the total differential of (ln) per-unit cost of housing is (see equation (6)):

$$\partial \ \text{ln} r_k = \frac{1}{1+\theta} \left(\partial \ \text{ln} \ R_k + \alpha \lambda^H_{kk|k} \partial \ \text{ln} \ W_k \right) \tag{B11}$$

Dividing by ∂g_k gives then equation (10).

Spillovers to other jurisdictions: Since in the model, changes in public good provision in one jurisdiction affect all other jurisdictions via in- and out-commuting, it is not easily possible to derive the total derivatives. To illustrate the difference between the partial derivatives derived above and the total derivatives, we focus in the following on the spillovers due to direct commuting. This means we focus on how changes in jurisdiction k affect jurisdiction s via out-commuting from k to s or in-commuting from s to k.

As in the paper, we start with the impact of out-commuting from k to s on jurisdiction s and how this may affect k again. Households of type ks live in k but work in s. Therefore, changes in households type ks, due to changes in public good provision in k, affect the number of workers in s and therefore wages in s. The size of the wage change in s is given by:

$$\partial \ \text{ln} \ w_s = \alpha \text{d} \ \text{ln} \ W_s = \frac{\alpha \lambda^W_{ks|s} \epsilon_g}{1 - \alpha(\epsilon_w - \epsilon_r \lambda^W_{ss|s} \lambda^H_{ss|s})} \partial \ \text{ln} \ v(g_k) \tag{B12}$$

This is very similar to the expression for the partial derivative for the wage in k; the only difference is that in s, it is not the workers that live in s that are affected but those that live in k. Incorporating the wage change in s (for all $s \neq k$) due to out-commuting from k in our partial derivatives for the number of residents in k gives:

$$\partial \ \text{ln} \ R_k = \epsilon_g \partial \ \text{ln} \ v(g_k) + \alpha \sum_{j=1}^{I} \left(\lambda^R_{kj|k} \epsilon_w - \lambda^H_{kj|k} \epsilon_r \right) \partial \ \text{ln} \ W_j \tag{B13}$$

This means the additional wage channel for the number of residents gets more important when taking the impact of out-commuting on out-commuters into account. However, in most cases the share of workers that live in the workplace jurisdiction ($\lambda^W_{kk|k}$) is larger than the share of in-commuters from one particular location ($\lambda^W_{ks|s}$), thus the main wage effect comes via workers than live in the workplace jurisdiction.

Since wage changes in s affect housing prices in s (and therefore the number of residents in s), this affects in-commuters from s to k (N_{sk}) as can be seen in equation (B7). Taking this additional channel into account means equation (B8) becomes:

$$\partial \ln W_k = \lambda^W_{kk|k}\epsilon_g \partial \ln v(g_k) + \alpha(\epsilon_w - \lambda^W_{kk|k}\epsilon_r\lambda^H_{kk|k})\partial \ln W_k - \alpha\epsilon_r \sum_{i \neq k} \lambda^W_{ik}\lambda^H_{ii|i}\partial \ln W_i \tag{B14}$$

Given that the wage changes in other jurisdictions are most likely modest, the resulting impact on in-commuters to k will be close to zero as the wage changes in s affect house prices in s only via workers in s that live in s.

We now turn to the impact of in-commuting from s to k. Our model incorporates that wage changes in k affect the number of in-commuters from other jurisdictions (N_{sk}), but it ignores the fact that changes in the wages in k affect the per-unit cost of housing in s. Incorporating these effects into the partial derivatives of (ln) number of workers gives:

$$\partial \ln W_k = \frac{\lambda^W_{kk|k}\epsilon_g}{1 - \alpha(\epsilon_w - \epsilon_r \sum_{i=1}^{I} \lambda^H_{ik|i}\lambda^W_{ik|k})} \partial \ln v(g_k) \tag{B15}$$

Thus, accounting for the impact of in-commuting on in-commuters means that the elasticity of workers with respect to wages is smaller, as the wage change is capitalized into house prices in all jurisdictions. The change in the elasticity is, however, modest at best since in most cases the share of workers that live in the workplace jurisdiction is larger than the number of in-commuters from a particular jurisdiction. Thus, the degree of capitalization of wage changes for out-commuters into house prices in their residence jurisdiction is likely to be small.

Since a house price change in s affects the number of residents and therefore the number of workers in s that live in s, accounting for in-commuting from other jurisdictions changes wages in s and therefore the number of out-commuters from k to s. The change in house price in s is $\lambda^H_{ks|s}\partial \ln v(g_k)$; this affects the number of workers in s via the share of workers in s that live in $s(\lambda^W_{ss|s})$ that respond by ϵ_r to the house price change. Incorporating this into the partial derivative of (ln) number of residents gives:

$$\partial \ln R_k = \epsilon_g \partial \ln v(g_k) + \alpha(\lambda^R_{kk|k}\epsilon_w - \epsilon_r\lambda^H_{kk|k})\partial \ln W_k + \alpha\epsilon_w \sum_{j \neq k}^{I} \lambda^R_{kj|k}\partial \ln W_j \tag{B16}$$

Again, this effect of in-commuting on out-commuters should be close to zero, as the wage effect in s is likely to be small, and house prices change only via workers in s that live in s.

Additionally, it is important to note that all additional effects on k due to spillovers into other jurisdictions are zero if wages are independent of the number of workers. Furthermore, it is important to note that changes in other jurisdictions affect not only k but all other jurisdictions as well. Given these considerations, we believe that the most important channels are captured by the partial derivatives that assume no changes in wages and per-unit housing costs in other jurisdictions.