I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16127

The Worth of Cities in Germany

Katja Gehr
Michael Pflüger

MAY 2023

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16127

The Worth of Cities in Germany

Katja Gehr
University of Würzburg

Michael Pflüger
University of Würzburg, CESifo and IZA

MAY 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Worth of Cities in Germany*

This paper extends the urban growth model of Duranton and Puga (2022) to explore the impact of cities on local firms and households and the aggregate economy of Germany. We adopt alternative micro-foundations for agglomeration economies and a non-linear specification of human capital accumulation. This allows us to characterize the social optimum and to bring the model in line with semi-endogenous growth. We also innovate by incorporating consumptive amenities and fiscal transfers into the model. On the empirical side we exploit the structural equations of the model and rich sets of micro-data for Germany's labor markets, housing rents, and household travel-to-work data, to estimate the population elasticities of urban benefits and costs. We are the first to establish elasticities for urban costs for Germany, an estimated elasticity of commuting costs with respect to distance travelled of 0.071, and an estimate for the population elasticity of travel congestion of 0.068. Our estimates for static and dynamic agglomeration elasticities are 0.017 and 0.020, respectively. We innovate on the calibration strategy to capture the important role of consumptive amenities and fiscal transfers in Germany. The model innovations and calibration are shown to be strongly supported by several pieces of evidence. Our key policy counterfactual is a proportionate increase of the population in Germany's Top Seven metropolises by 10% which implies a significant overall welfare benefit of 1.12% per person. This involves mild losses for city incumbents but strong gains for city newcomers. We also address the effects of a counter-factual shift to the social optimum and a counterfactual removal of fiscal transfers. Our final exercise evaluates the implications of cities and agglomeration economies for aggregate growth in Germany. We find that these account for 0.011 additional percentage points of growth in income per capita per year.

JEL Classification: C52, R12, D24

Keywords: city systems, urban growth, agglomeration and dispersion economies, land-use regulations

Corresponding author:
Michael Pflüger
Faculty of Economics
University of Würzburg
Sanderring 2
97070 Würzburg
Germany

E-mail: michael.pflueger@uni-wuerzburg.de

* We thank Gilles Duranton for very insightful comments and discussion. We also acknowledge helpful comments and suggestions from Rainald Borck, Toker Doganoglu, Thomas Klier, Eric Mayer, Kurt Schmidheiny and audiences in Essen, Montreal, Passau and Würzburg. We would like to express particular thanks to Tobias Seidel, Marcel Henkel and Jens Suedekum for providing us their German tax-transfer data.

1 Introduction

What is the worth of cities in Germany? This paper draws and innovates on recent advances in the theory and empirics addressing urban systems and urban growth to explore the role of cities for local firms and households and for Germany's aggregate economy.

Cities have for long been hailed as engines for prosperity and growth (Jacobs 1969, 1984; Lucas 1988; Glaeser 2011). Their evolution has recently produced much discontent, however. Citizens and would-be city dwellers are engaged in urban battles over space, high rents and house prices in places that they deem attractive. The Top Seven metropolises in Germany, Berlin, Hamburg, Munich, Cologne, Frankfurt, Stuttgart and Duesseldorf, are a focal point of public attention.[1] These issues also prevail in other large cities in Germany and around the world.[2] Panel A of Figure 1 shows that Germany's large cities pay high wages, one indicator for their attractiveness, whilst panel B depicts a key factor for urban discontent, high rents in these locations.[3]

Panel A: Mean annual earnings 2017 Panel B: Rents per sqm 2017

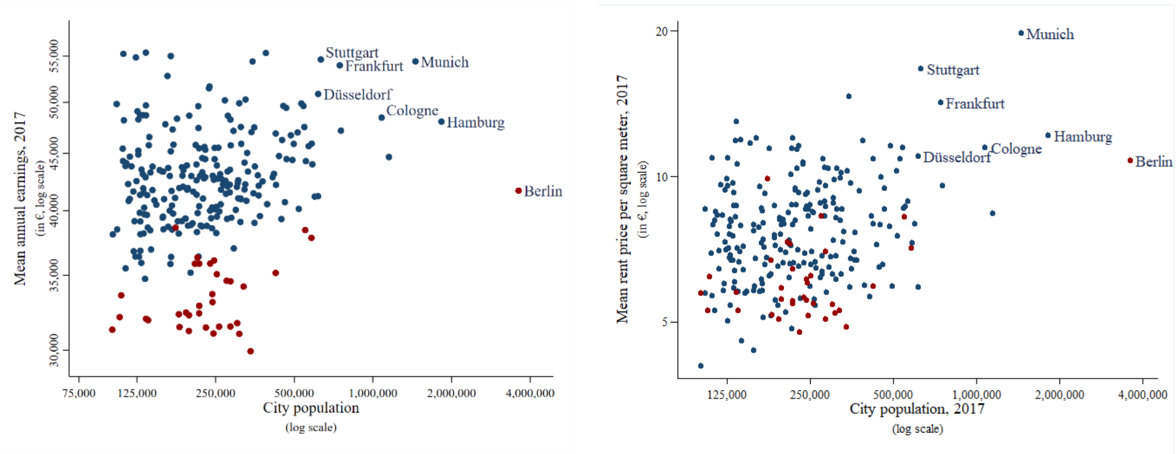

Figure 1: Mean annual earnings 2017 and rents per sqm 2017 in German Cities

Notes: Panel A links mean annual earnings to the city population in 2017. Panel B plots the relationship between the mean rent price per sqm in a city and its city population in 2017. Cities are defined as in appendix A1. Blue/red dots indicate cities in West/East Germany. The Top Seven are highlighted by name. See section 3.1 and Appendix A for the data.

Recent research, largely directed at the United States, puts various urban containment policies, land-use regulations, housing supply constraints and public frictions (e.g. fiscal externalities and infrastructure policies) at center stage to rationalize these urban battles and to gauge their

[1] The mentioned metropolises, termed the "Top 7", "Big 7", or "A-locations", are much-hyped by the media and the real estate industry (Economist 2015; Postbank 2022; PricewaterhouseCoopers 2019).

[2] Silicon Valley is the crystallization point in the United States (Economist 2018).

[3] Notice that these patterns hold true both in the West (black dots) and the East (blue dots), as most clearly documented (but not marked in the figure) by the two biggest cities in Germany's East after Berlin, Leipzig (~ 580,000 citizens) and Dresden (~ 550,000 citizens).

local and aggregate effects.[4] Further research highlights that these restrictions result from a political game between city incumbents and local policymakers, who strive to contain city access through various forms of 'NIMBYism', and would-be citizens, who desperately look for 'affordable housing' (Glaeser 2020; Glaeser and Cutler 2021). Duranton and Puga (2022), henceforth DP (2022), pull these research strands together to analyze the US city system. They propose an urban growth model with local governments choosing city sizes which maximize the welfare of city incumbents. Locally optimal sizes are shored up by land-use regulations imposed on would-be citizens. Their model also features fundamental city drivers grounded in micro-founded benefits and costs of cities and human capital accumulation, as well as a random growth component.

We innovate in three ways on the DP (2022) model to address Germany's city system. First, we adopt alternative micro-foundations of agglomeration economies based on input sharing with monopolistically competitive local suppliers following Duranton and Puga (2004; 2014), whilst DP (2022) assume input sharing with competitive suppliers and human capital spillovers.[5] Second, we modify human capital accumulation such that it exhibits diminishing returns and thus accords with semi-endogenous growth (Jones 2002; 2005; 2022). These two modifications have important payoffs. First, they allow us to establish the social optimum of the model.[6] Second, by eliminating linearity in human capital accumulation (implied in DP 2022), we free the model from a permanent growth driver which would imply ever fewer and bigger cities. Our third theoretical innovation concerns the incorporation of consumptive amenities and fiscal transfers. Whilst the US city system can be well-approximated empirically by focusing on differences in productive amenities (DP 2022), the same is not true for Germany. Rather, in Germany, as in much of Europe, consumptive amenities are key for cities (Brueckner et al. 1999). Berlin, with 3.6 million citizens, by far the largest city in Germany, is a case in point. The mean wage in Germany's capital is 18%-30% lower than in Hamburg (1.8 million), Munich (1.46 million), Frankfurt (747.000), and many smaller-sized cities, suggesting that

---

[4] See Davis et al. (2014), Desmet and Rossi-Hansberg (2013), Fajgelbaum et al. (2019), Fajgelbaum and Gaubert (2020), Hsieh and Moretti (2019), Turner et al. (2014) and the surveys by Proost and Thisse (2019) and Duranton and Puga (2020). Henkel et al. (2021) is one exception focusing on the tax-transfer system in Germany.

[5] The micro-foundations of agglomeration economies in Duranton and Puga (2014; 2004) draw on Ethier (1982) and Abdel-Rahman and Fujita (1990) but, unlike the latter, abstain from assuming that labor is needed in addition to intermediates in the final production stage. Under the assumption that the total differential land rent is proportionally rebated to citizens, agglomeration economies exhibit no net inefficiency, because the prices of all inputs reflect an identical mark-up on marginal costs (in contrast to the Abdel-Rahman and Fujita (1989) model which features a distortion), see Duranton and Puga (2004, section 2.24) and Pflüger (2021). This property is important when we establish the social optimum of our model, see below.

[6] DP (2022) do not address the social optimum of their model, they only argue that the private and social return to the marginal resident is not equated across cities, so that aggregate consumption is inefficient (DP 2022:11).

Berlin's attractiveness stems from its consumptive amenities rather than from an overwhelming productive capacity.[7] Moreover, we take into account that a powerful fiscal equalization system is in place in Germany, a complicated web of transfers across different layers of government mainly targeted at regional redistribution, which has been shown to be an important location determinant (Henkel et al. 2021). We treat these fiscal transfers as exogenous factors.

On the empirical side we build on the structural equations of the model to obtain estimates for the elasticities of urban benefits and costs. We are, to the best of our knowledge, the first to establish elasticities for urban costs for Germany. We draw on rich sets of micro-data for the German labor market (IAB's SIAB-R, see Antoni et al. 2019; Dauth and Eppelsheimer 2020; Eberle and Schmucker 2019), for housing rents (RWI's GEO-RED, see RWI 2019; Schaffner 2020), and for household travel-to-work data within cities (BMVI 2017). Our study is the first which quantifies and uses geographical constraints for Germany's cities, following the methods established by Saiz (2010). We estimate highly significant static and dynamic agglomeration elasticities of 0.017 and 0.020, respectively. Our estimated elasticity of commuting costs with respect to distance travelled is 0.071, and our estimate for the population elasticity of travel congestion is 0.068, both are highly significant.

Equipped with these estimates and a parameter for the population elasticity of rural income from the extant literature, we map the German economy into the general equilibrium with local governments. Our calibration of the model differs significantly from DP (2022) who back out productive amenities from the observed US city sizes by inverting the model's equation for city sizes under local governments. This turns out to yield very plausible results for the United States, where New York City is the biggest city and hence, also implied to be the most productive city (and so on for the other cities). This calibration procedure would not deliver meaningful results for the German city system, however. This can be rationalized by reference to Berlin, for the reasons spelled out above. Our strategy is therefore to use observed German wages and city population sizes, and the model-implied wage equation to back out productive amenities in a first step. The population size of Berlin (and that of all other German cities) is then rationalized in a second step by consumptive amenities, which we back out from the model-implied locally optimal city sizes under local governments, given the observed city populations and the productive amenities derived in the first step. We then show that the resulting model and calibration are strongly supported by several pieces of independent

[7] This accords well with the famous words of former mayor Klaus Wowereit, who characterized Berlin as "Poor, but sexy" (Economist 2017).

evidence. The evolution of college-educated population shares in Germany exhibits different levels but a parallel trend across broad size classes of cities, thus lending support to the model's human capital formation process. The model-implied real income ranking of Germany's cities is strongly validated by a striking correlation with an independent measure of real income based on different data. The housed price gradients and, notably, house prices at the edges of German cities, accord with the land-use regulations implied by the political-economy mechanism of the model and so do house prices in the city centers. These four facts make us confident about our model innovations and calibration.

Our key policy counterfactual is inspired by extant research on the United States which focuses on removing mobility restrictions in key agglomerations.[8] Given the paramount public attention dedicated to Germany's Top Seven, we focus on a proportionate increase of the population of 10% in these cities. This counterfactual, implemented by corresponding reductions in land-use regulations, implies a decrease of the population in rural areas from initially 9 million to 8.3 million people. One city would be vacated, so that 263 cities remain. The overall (national) welfare effect is a real income benefit of 1.12% per person. Incumbents in the top seven cities loose, but their real income decreases by a mild 0.002% only, whilst city newcomers, the rural population, and the incumbents of the vacated city experience a strong increase of 1.26% per person. This indicates that urban containment policies in the Top Seven in Germany have significant societal costs. We also address a counterfactual shift to the social optimum which would imply yet stronger population movements and welfare gains and a counterfactual removal of all fiscal transfers across cities which would deliver an aggregate welfare gain. Our final exercise is an evaluation of the implications of cities and agglomeration economies for aggregate growth in Germany. We find that these account for 0.011 additional percentage points of growth in income per capita per year.

Related Research. Our paper is most closely related to the model and methodology proposed in DP (2022). However, we undertake significant innovations to address the German city system. These range from the model specification (micro-foundation of agglomeration economies, process of human capital accumulation, incorporation of consumptive amenities and fiscal transfers) to our entirely different model calibration. The DP (2022) model itself builds on the extension of the static city system model of Henderson (1974) considered in Albouy et al. (2019) and in Behrens and Robert-Nicoud (2015), the urban growth model of

[8] Hsieh and Moretti (2019) focus on New York and the San Francisco Bay Area. Duranton and Puga (2022) address New York, Los Angeles, San Francisco - Oakland - San Jose, Washington DC, Boston, Seattle, and San Diego.

Black and Henderson (1999), and related work (Rossi-Hansberg and Wright 2007; Duranton 2007), and bridges it with research on neoclassical growth (Lucas 1988) and random growth (Gabaix 1999; Eeckhout 2004). Our specification connects the model with current research on semi-endogenous growth (Jones 2005; 2022).

A burgeoning literature on new quantitative spatial modelling offers a related, yet different theoretical backbone for quantitative analysis (Ahlfeldt et al. 2015; Allen and Arkolakis 2014; Redding and Rossi-Hansberg 2017; Redding 2020) which allows for trade costs across locations, and hence, home market effects of the new economic geography which are absent in the model of DP (2022). Henkel et al. (2019) provide an important contribution along these lines which addresses the welfare and location effects of Germany's system of fiscal transfers which we take up below. The DP (2022) model comes with the advantages to bring out the fundamental trade-off between agglomeration and dispersion of urban economics (Fujita and Thisse 2013), to propose an alternative political-economy-based mechanism of city size determination, and to tie up the analysis with urban growth, taking both fundamental drivers and a random growth component into account.

Our paper is also related to various empirical literatures. We already referred to a literature addressing various urban containment policies within different models (cf. footnote 4). Second is research concerned with the estimation of the costs and benefits of cities (e.g. Ahlfeldt and Pietrostefani 2019; Combes and Gobillon 2015; Combes et al. 2011; Combes et al. 2019; and the survey by Duranton and Puga 2020). In the German context, agglomeration economies have recently been estimated by Grujovic (2021) and Dauth et al. (2021) yielding similar results as ours. To the best of our knowledge, our analysis is the first which provides estimates of congestion elasticities in cities for Germany. Our research is finally also connected to recent research addressing local price level indices and real wage inequality across locations which has a strong focus on skill group differences (e.g. Moretti 2013; Diamond 2016; Diamond and Moretti 2021; Diamond and Gaubert 2022; Dustmann et al. 2022; Weinand and von Auer 2020), whilst we are interested in the real income situation of city incumbents and outsiders.

The structure of our paper is as follows. Section 2 presents the model, derives the general equilibrium under the local government allocation and the social optimum. Section 3 describes the data, the empirical estimation of the agglomeration and congestion elasticities, and the quantification of the model. Section 4 takes up the counterfactuals and derives the effects on national welfare, the distributional effects, and the growth effects. Section 5 concludes.

2 The model

2.1 Economic environment

Our theoretical framework draws on the urban growth model developed by DP (2022), which we modify by building on different micro-foundations, a non-linear specification of human capital accumulation, and by incorporating consumptive amenities and fiscal transfers.

Space and population. The model is set in discrete time t. There is a fixed amount of land which can be occupied by cities (i) or used as rural area (r) for rural production. We follow DP (2022) in assuming that land sites are heterogeneous in terms of their time varying natural production amenities A_{it}^{p} and in terms of their time-invariant geographical constraints to city development. We additionally assume that land sites differ in terms of their time-invariant consumption amenities A_{i}^{c}. The economy hosts an exogenously evolving population N_t which resides in cities and the rural area. Both the extensive margin of urban development (the number of sites to be developed as cities) and the intensive margin (the population in each city N_{it} and the rural area N_{rt}) are determined endogenously. The allocation is governed by local city governments which perform three actions. First, they rent land at its alternative cost and sublease plots for one period to the highest bidder at each location. Second, they redistribute the resulting differential land rent among the local populations. Third, and crucially, they choose the city size ('housing stock') to maximize the welfare of city residents, and they control the locally optimal size by imposing land-use regulations on would-be city entrants.

The population is replaced each time-period. At the end of a period each person has one offspring and then dies. The offspring 'inherits' the location of the parent and the economy's average education level. Production amenities in each site are governed by a multiplicative process involving non-negative shocks g_{it} which are independently drawn from a common distribution with support $(1, \infty)$, such that $A_{it}^{p} = g_{it} A_{it-1}^{p}$. Faced with the updated productivity amenity, individuals decide whether to remain at the inherited location or to move elsewhere, facing the respective land rents and land-use regulations.

The economy produces one homogeneous final output with one production factor, educated labor ('human capital'), both in the cities (Y_{it}), and in the rural area (Y_{rt}). This output is the numéraire and assumed to be tradeable at no cost. Individuals' utility increases in the consumption of the numéraire. Each individual supplies 1 unit of labor without loss of utility. A city dweller incurs urban costs for commuting to the workplace at the central business district

(CBD) and because of the housing costs for 1 unit of land/floorspace which she inelastically demands. Individuals in the rural area are assumed to face no urban costs.

Human capital formation. Human capital H_{it} in a city is related to its workforce L_{it} through the identity $H_{it} = h_{it}\, L_{it}$, where h_{it} is the level of human capital per worker. The accumulation of human capital is specified by $h_{it} = h_t \cdot {N_{it}}^{\beta}$. The first component, $h_t = e^{\psi_t\, l_{ht}}$, $\psi_t > 1$, is a Mincer-process as in Jones (2002; 2005), where $0 < l_{ht} < 1$ denotes the share of each individual's unit time endowment devoted to schooling.[9] The second component, ${N_{it}}^{\beta}$, $\beta > 0$, is a city-specific learning effect which is the stronger, the bigger the city, in accordance with empirical findings (De la Roca and Puga 2017). The local workforce is then $L_{it} = (1 - l_{ht}\,)N_{it}$ and a city's human capital is $H_{it} = B(l_{ht})\, {N_{it}}^{1+\beta}$, where $B(l_{ht}) \equiv\ e^{\psi_t\, l_{ht}}(1 - l_{ht}\,)$. We note that $B(0) = 1$ and $B'\ (l_{ht}) > 0$.

Production in cities. Production of the numéraire in city i is modelled as in Duranton and Puga (2004; 2014), cf. footnote 5. The final good is produced by combining an (endogenous) mass m_{it} of symmetric intermediate goods ω with quantities $q_{it}(\omega)$ according to a CES production function $Y_{it} = A_{it}^{p}\left\{\int_0^{m_{it}} q_{it}(\omega)^{\frac{1}{1+\sigma}}\, d\,\omega\right\}^{1+\sigma}$, where $0 < \sigma < 1$. The elasticity of substitution between any two intermediates is $\varepsilon \equiv (1+\sigma)/\sigma$ and A_{it}^{p} represents the local production amenity. Intermediates are non-tradable and produced under increasing returns and monopolistic competition with educated labor $h_{it}(\omega)$ according to the function $h_{it}(\omega) = \frac{\alpha}{\rho} + \frac{q_{it}(\omega)}{\rho}$ which exhibits a fixed and a variable (output-related) component. The quantities of intermediates are chosen to minimize the costs to produce final output. Conditional input demand is $q_{it}(\omega) = \frac{[z_{it}(\omega)]^{-(1+\sigma)/\sigma}}{\left\{\int_0^{m_i}[z_{it}(\omega')]^{-1/\sigma}d\omega'\right\}^{1+\sigma}}\,\frac{Y_{it}}{A_{it}^{p}}$, where $z_{it}(\omega)$ denotes the price of intermediate ω. This implies that firm ω faces an own-price demand elasticity of $-(1+\sigma)/\sigma$. Its profit-maximizing price is a constant mark-up on marginal costs, $z_{it}(\omega) = (1+\sigma)\, w_{it}$, where w_{it} is the wage paid per unit of educated labor. Symmetry allows us to drop the index ω from now on. Free entry drives intermediates' profits to zero, $\pi_i = z_{it}q_{it} - w_{it}h_{it} = 0$. A firm's break-even output is then $q_{it} = \frac{\alpha}{\sigma}$, and its demand for educated labor is $h_{it} = \frac{\alpha(1+\sigma)}{\rho\sigma}$. Labor market clearing implies that the mass of intermediates is $m_i = \frac{H_i}{h_i} = \frac{\rho\sigma}{\alpha(1+\sigma)}H_i$. Invoking these results

[9] The exponential formulation is possibly the simplest and most straightforward way to incorporate human capital such that it accords with the research on schooling and wages (see Jones 2002 and also Bils and Klenow 2000).

and symmetry, final output in city i is $Y_i = A_{it}^p \, H_i^{1+\sigma}$, where the choice $\rho = (1+\sigma)(\alpha/\sigma)^{\frac{\sigma}{1+\sigma}}$ normalizes units of intermediates (as in Duranton and Puga 2014). Combining this with human capital in the city we have $Y_i = A_{it}^p \, B(l_{ht})^{1+\sigma} \, N_{it}^{(1+\beta)(1+\sigma)}$.

Educated workers are compensated with their average product:

$$w_{it} = \frac{Y_{it}}{N_{it}} = A_{it}^p \, B(l_{ht})^{1+\sigma} \, N_{it}^{\sigma+\eta}, \quad \eta \equiv \beta(1+\sigma) \tag{1}$$

Eq. (1) shows that wages grow with city size and the premium associated with larger cities has two sources. The parameter σ implies a static earnings premium from working in a bigger city which stems from the productive advantages of sharing local intermediates. The parameter η captures a dynamic earnings premium associated with (the parameter of) learning in cities β which interacts with the static agglomeration externality.[10]

Production and utility in the rural area. Production of the numéraire in the rural area takes place with a decreasing marginal product of labor earning workers a wage w_{rt} as specified in $w_{rt} = A_{rt} \, N_{rt}^{-\lambda}$, where $0 < \lambda < 1$, and where all location attributes are captured in A_{rt}.[11] Since this marginal product approaches infinity as rural employment N_{rt} approaches nil, rural production is always active in the economy. Individuals in the rural area are assumed to face no urban costs. We proceed as follows to integrate the status quo fiscal transfers in Germany which we take as exogenously determined. These transfers redistribute income across locations and exhibit the consistency requirement to sum up to zero (Henkel et al. 2021). To capture the consequences for the rural area, we sum up the transfers across all city locations. The (negative) balance which we denote by the constant K_r then captures the transfers from cities to the rural area. We assume K_r to be distributed on a per capita basis implying a transfer of κ_r per person in the rural area. The utility of rural workers is then given by (a straightforward extension of Albouy et al. 2019 and DP 2022):

$$v_{rt} = w_{rt} + \kappa_r = A_{rt} \, N_{rt}^{-\lambda} + \kappa_r \tag{2}$$

[10] This differs from DP (2022) where the dynamic component is represented by the exponent β. The reason for this difference is not the different micro-foundation of agglomeration economies. Rather, it is DP's (2022) assumption of an asymmetry between the supply of local human capital available for production (where learning in cities is taken into account) and the supply of local human capital in idea generation (where the effect of learning in the city is ignored), see DP (2022, FN 9). If the effect of learning in cities was accounted for in both, the dynamic component in DP (2022) would also be given by $\eta \equiv \beta(1+\sigma)$.

[11] To keep the model simple and tractable we follow DP's (2022) implicit assumption that education does not raise a worker's marginal product in rural production.

Indirect utility in cities. Indirect utility a citizen $v_{it}(w_{it}, nuc_{it}, A_i^c, \varphi_i)$ is positively affected by her wage w_{it}, negatively affected by net urban costs nuc_{it} (specified below), positively affected by time-invariant consumption amenities A_i^c , and positively (negatively) affected by fiscal transfers φ_i to city locations if positive (negative). Our specification of indirect utility is

$$v_{it}(w_{it}, nuc_{it}, A_i^c, \varphi_i) = w_{it} - \frac{(1-\varphi_i)}{A_i^c} nuc_{it} \quad (3)$$

which has two key properties. First, it keeps up the gist and tractability of the DP (2022) model. Second, as shown below, under this specification consumptive amenities and fiscal transfers affect not only utility but also city sizes in meaningful ways.[12] Since all location factors are incurred in the numéraire, indirect utility eq. (3) can also be understood as net-consumption or real income in terms of the final good. We will use these terms interchangeably from now on. A very similar utility function is used in recent extensions of the Rosen-Roback model (Diamond and Gaubert 2022; Kline and Moretti 2014; Moretti 2011)[13]. Our incorporation of fiscal transfers is similar as the modelling of fiscal externalities in Albouy et al. (2019).

Internal structure of cities. Cities are monocentric, linear one-sided, and stretch from the CBD at $x = 0$ to the city border x_b. Due to time-invariant geographical constraints only the share $0 < \Lambda_i < 1$ of the raw land of a site can be developed (as in Brueckner 1987 and Saiz 2010). Each worker consumes 1 unit of land/floor-space and, hence, $1/\Lambda_i$ units of raw land.[14] The city border is then at $x_b = N_{it}/\Lambda_i$. Commuting costs in the city are formalized by $T_{it}(x) = \tau_t \, N_{it}{}^{\theta} x^{\gamma}$, where x^{γ} is the length of the commute which is assumed to increase with elasticity $\gamma > 0$ with distance x from the CBD to reflect geographical obstacles. The term $\tau_t \, N_{it}{}^{\theta}$ formalizes the cost per unit distance, where τ_t is a parameter for the commuting technology, and $N_{it}{}^{\theta}$ stands for congestion in commuting in the city which relates to city size N_{it} with elasticity $\theta > 0$. Spatial equilibrium in the city commands that urban costs (which comprise rents and commuting costs) are equalized across all locations x so that,

$$T_{it}(x) + \frac{1}{\Lambda_i} R_{it}(x) = \frac{1}{\Lambda_i} R_{it}(0) = T_{it}(N_{it}/\Lambda_i) \quad (4)$$

[12] City size would be unaffected if consumptive amenities were included as multiplicative shifters of per capita income net of urban costs. Our specification follows the rationale that guides Duranton and Puga (2022) to model commuting costs in terms of local output rather than in terms of time, because city size is unaffected by productive amenities under the latter, see Duranton and Puga (2004) and Albouy et al. (2019).

[13] These works typically also impose separability assumptions on wages, housings costs, consumptive amenities and taxes, for convenience. They typically allow for an idiosyncratic location preference factor.

[14] We follow DP (2022) in abstracting from a competitive construction industry. Hence, there are no other costs than land to provide ("build") new homes, and we can therefore equate the terms land and floorspace.

where $R_{it}(x)$ is the bid rent per unit of raw land at location x. Define $P_{it}(x) \equiv R_{it}(x)/\Lambda_i$ to be the (quality-transfer) unadjusted price of a dwelling at distance x from the CBD.[15] Commuting costs from the city border to the CBD, $T_{it}(N_{it}/\Lambda_i)$, and the price of a dwelling at the CBD, $P_{it}(0)$, are compact measures for equilibrium urban costs. Eq. (4) implies the familiar result that, in a spatial equilibrium in the city, an increase in commuting costs associated with a longer commute must be compensated by a corresponding fall in land rents, $d\,P_{it}(x)/dx = -\,d\,T_{it}(x)/dx$, the so-called Alonso-Muth condition. Using the formula for commuting costs, $T_{it}(x) = \tau_t\,{N_{it}}^{\theta} x^{\gamma}$, bid rents for raw land at x implied by (4) are calculated as $R_{it}(x) = \Lambda_i\,\tau_t\,{N_{it}}^{\theta}((N_{it}/\Lambda_i)^{\gamma} - x^{\gamma})$. Total differential land rents in the city, $TLR_{it} = \int_0^{x_b} R_{it}(x)\,dx = {N_{it}}^{\gamma+\theta+1}\,\tau_t\,\gamma/(1+\gamma){\Lambda_i}^{\gamma}$, are assumed to be rebated to citizens on a per capita basis. Net urban costs per capita in the city are then given by:

$$nuc_{it} = \left[T_{it}(x) + \frac{1}{\Lambda_i}\,R_{it}(x)\right] - \frac{TLR_{it}}{N_{it}} = \frac{\tau_t}{(1+\gamma){\Lambda_i}^{\gamma}}{N_{it}}^{\gamma+\theta} \tag{5}$$

Using eqs. (1) and (5) in eq. (3) indirect utility can be written as:

$$v_{it} = A_{it}^{p}\,B(l_{ht})^{1+\sigma}\,{N_{it}}^{\sigma+\eta} - \frac{(1-\varphi_i)}{A_i^c{\Lambda_i}^{\gamma}}\frac{\tau_t}{(1+\gamma)}{N_{it}}^{\gamma+\theta} \tag{6}$$

2.2 Equilibrium allocation with local governments

Human capital investment. Citizens choose the time they devote to learning, l_{ht}, to maximize indirect utility eq. (6). This problem boils down to maximizing $B(l_{ht}) = e^{\psi_t\,l_{ht}}(1-l_{ht})$. We use an asterisk (*) to denote the local government solution from now on. The solution is $l_{ht}^{*} = 1 - 1/\psi_t$ (as in Jones 2005) and we define $B(l_{ht}^{*}) = e^{\psi_t\,l_{ht}^{*}}(1-l_{ht}^{*}) = e^{\psi_t - 1}/\psi_t \equiv B(\psi_t)$.

Determination of city size by incumbents. City size is determined by incumbent residents maximizing their indirect utility. Using $B(\psi_t)$ in eq. (6) indirect utility can be written as:

$$v_{it} = A_{it}^{p}\,B(\psi_t)^{1+\sigma}\,{N_{it}}^{\sigma+\eta} - \frac{(1-\varphi_i)}{A_i^c{\Lambda_i}^{\gamma}}\frac{\tau_t}{(1+\gamma)}{N_{it}}^{\gamma+\theta} \tag{7}$$

'Locally optimal' city sizes and the associated indirect utility levels are then[16]

$$N_{it}^{*} = \left(\frac{(\sigma+\eta)(1+\gamma)}{(\gamma+\theta)}\frac{B(\psi_t)^{1+\sigma}}{\tau_t}\frac{A_{it}^{p}\,A_i^{c}\,{\Lambda_i}^{\gamma}}{(1-\varphi_i)}\right)^{\frac{1}{(\gamma+\theta)-(\sigma+\eta)}} \tag{8}$$

[15] In view of eq. (3), the spatial equilibrium condition (4) could also be written by multiplying each term with the factor $(1-\varphi_i)/A_i^c$. This quality-transfer adjustment is inessential for the trade-off between rents and commuting costs within a city. However, it is relevant when taking the model to the data.

[16] The second-order condition is $\gamma + \theta > \sigma + \eta$. Positive city sizes require agglomeration economies, $\sigma + \eta > 0$.

$$v_{it}^{*} = \frac{(\gamma+\theta)-(\sigma+\eta)}{(\sigma+\eta)(1+\gamma)} \frac{\tau_t\,(1-\varphi_i)}{A_i^c\,\Lambda_i{}^{\gamma}} N_{it}^{*}{}^{\gamma+\theta} \qquad \eta \equiv \beta(1+\sigma) \tag{9}$$

Eq. (8) reflects the "fundamental trade-off" between agglomeration and congestion forces in cities (Fujita and Thisse 2013) which selects the peak of the hump-shaped real income curves implied by eq. (7). As in DP (2022), city sizes are positively related to the level of productive amenities A_{it}^{p} and the share of locally developable land Λ_i, and also positively affected by human capital accumulation which positively responds to the return to education. In addition, city sizes are positively affected by the level of consumptive amenities A_i^c, and positively (negatively) affected by fiscal transfers φ_i if positive (negative). Eq. (9) shows that (locally optimal) real income v_{it}^{*} rises in (locally optimal) city size N_{it}^{*} with elasticity $\gamma+\theta$. By implication, land sites will be occupied by cities in descending order of the location-specific composite factor $\Omega_{it} \equiv A_{it}^{p\,\frac{\gamma+\theta}{(\gamma+\theta)-(\sigma+\eta)}} \left(\frac{A_i^c\,\Lambda_i{}^{\gamma}}{(1-\varphi_i)}\right)^{\frac{\sigma+\eta}{(\gamma+\theta)-(\sigma+\eta)}}$, from the highest to successively lower ones. Note that, using $P_{it}(0) = \frac{\tau_t}{\Lambda_i{}^{\gamma}} N_{it}{}^{\gamma+\theta}$, we can rewrite eq. (9) as:

$$v_{it}^{*} = \frac{(\gamma+\theta)-(\sigma+\eta)}{(\sigma+\eta)(1+\gamma)} \frac{(1-\varphi_i)}{A_i^c} P_{it}(0) \tag{10}$$

This reveals that the (quality-transfer) adjusted price of housing in the city center, $\frac{(1-\varphi_i)}{A_i^c} P_{it}(0)$, is a sufficient statistic for the attractiveness of a site (see also footnote 15). Hence, our model extension exhibits the property that the maximization of the value of individual homes, Fischel's (2001) homevoter hypothesis, is equivalent to the maximization of v_{it}. DP (2022) term this the "golden rule of planning regulation". This theoretical implication is strongly supported by the German data, as we show below. We follow DP (2022) in assuming that these real income differences across locations are shored up by land-use regulations p_{it} which are imposed by city incumbents on (potential) city newcomers to make them indifferent between living in a city and living in the rural area where the lowest level of real income $v_{rt}^{*} = w_{rt} + \kappa_r$ is realized in this economy. The level of land-use regulations that newcomers face in city i is thus given by

$$p_{it}^{*} = v_{it}^{*} - v_{rt}^{*}. \tag{11}$$

The city system in general equilibrium. The general equilibrium of the city system is characterized by (locally optimal) city sizes (8), real income levels in cities (9), real income in the rural area given by $v_{rt}^{*} = A_{rt}\, N_{rt}^{*}{}^{-\lambda} + \kappa_r$, and by local land-use regulations (11). The model is closed by two conditions. First, real income in the marginally populated city (characterized by an underline) just equals the real income in the rural sector, $\underline{v}_{it}^* = v_{rt}^*$, so that this marginal city does not have to impose land-regulations, i.e. $\underline{p}_{it}^* = 0$. Second, the population in the cities and in the rural area must add up to the total population N_t.

City growth and the size distribution of cities. We now turn to the determinants of city growth and the city size distribution implied by the model. Log-differencing eq. (8) yields a decomposition of the factors that determine the growth of a city i over two consecutive points in time $t-1$ and t:

$$\Delta \ln N_{it} \equiv \ln N_{it} - \ln N_{it-1} = \frac{1}{(\gamma+\theta)-(\sigma+\eta)} \left[\Delta \ln A_{it}^p + (1+\sigma)\Delta \ln B(\psi_t) - \Delta \ln \tau\right] \quad (12)$$

Starting from the right end of eq. (12), a first systematic component of city growth is given by the evolution of the common commuting technology $\Delta \ln \tau$. A second systematic component arises from human capital accumulation through time spent learning which enhances human capital. This component exerts a positive influence on city sizes if the return to education increases. The third component, $\Delta \ln A_{it}^p$, is used by DP (2022) to connect this city systems model with random growth models (see Gabaix 1999, Eeckhout 2004; Duranton and Puga 2014). Under the assumption that local production amenities accumulate through non-negative random multiplicative shocks g_{it} which are identically and independently distributed across locations, $\Delta \ln A_{it}^p = \ln g_{it}$, it follows from eq. (12) that cities have a common growth trend in expectation, but they also experience idiosyncratic ups and downs relative to this trend. The growth process described in eq. (12) thus follows Gibrat's law so that (under some further conditions spelled out in DP 2022), the steady-state city size distribution implied by the model approximates Zipf's law, a very desirable feature of the DP (2022) model in view of the empirics in many countries, including Germany.[17]

2.3 The social optimal allocation

The social optimal allocation maximizes aggregate utility. Under the assumptions this is equivalent to the maximization of aggregate net consumption of the numéraire subject to the population constraint (as in Albouy et al. 2019).[18] Since our focus is not on the normative issue of the optimality of the German fiscal transfer system, we only include transfers for their positive effects and treat them as exogenously given (cf. section 1). In this sense, the social

[17] Duranton and Puga (2014) survey these studies and argue forcefully that looking at city-size distributions through the lens of Zipf's law is useful, despite the finding that a (truncated) log normal or a double Pareto log normal distribution provides yet a better fit for some countries, see e.g. Eeckhout (2004) and Giesen et al. (2010).
[18] It is important to note here that our micro-foundations of agglomeration economies exhibit no inefficiency, on net, cf. footnote 5. This property is implicitly maintained by Albouy et al. (2019:104).

optimum characterized in the following is a constrained one.[19] The program is to choose $\{l_{ht}, N_{it}, N_{rt}, \underline{A}_{it}^{p}\}$, i.e. the time devoted to learning, city populations, the rural population, the productive amenity of the marginal city, to maximize:

$$Y_{rt}(N_{rt}) + \int_{A_{min}^{p}}^{\infty} N_{it}\, v_{it}\, dG(A_{it}^{p}) \tag{13}$$

$$\text{s.t.} \qquad N_{rt} + \int_{A_{min}^{p}}^{\infty} N_{it}\, dG(A_{it}^{p}) = N_t \tag{14}$$

where $Y_{rt}(N_{rt}) = \frac{A_{rt}}{1-\lambda} N_{rt}{}^{1-\lambda} + \kappa_r N_{rt}$ is rural real income and where $G(A_{it}^{p})$ is the cumulative density of the productive amenity in the economy with support $[A_{min}^{p}, \infty]$. Using μ to denote the Lagrange-Parameter associated with the population constraint (14) and the index 'opt' to characterize the social optimum, the first order conditions of this program imply:

$$l_{ht}^{opt} = 1 - \frac{1}{\psi_t} \tag{15}$$

$$\mu = Y_{rt}{}'(N_{rt}^{opt}) = A_{rt} N_{rt}^{opt^{-\lambda}} + \kappa_r \tag{16}$$

$$\mu = (1+\sigma+\eta)\, A_{it}^{p} \cdot B(\psi)^{1+\sigma} \cdot N_{it}^{opt^{\sigma+\eta}} - (1+\gamma+\theta)\, \frac{(1-\varphi_i)}{A_i^c \Lambda_i{}^{\gamma}} \frac{\tau_t}{(1+\gamma)}\, N_{it}^{opt^{\gamma+\theta}} \tag{17}$$

$$\mu = \underline{A}_{it}^{p\,opt}\, B(\psi)^{1+\sigma}\, \underline{N}_{it}^{opt^{\sigma+\eta}} - \frac{(1-\varphi_i)}{A_i^c \Lambda_i{}^{\gamma}} \frac{\tau_t}{(1+\gamma)}\, \underline{N}_{it}^{opt^{\gamma+\theta}} \tag{18}$$

together with the population constraint (14). Eq. (15) gives the socially optimal share of time devoted to education. This choice is undistorted and so the resulting optimal level coincides with the privately optimal level derived in the previous section (as in Jones 2005). Results (16) - (18) mirror the results of Albouy et al. (2019). Eq. (16) is the extensive margin condition for development which states that the shadow price of any further worker corresponds to the marginal product of labor (plus the fiscal balance) in the rural area. Eq. (17) is the intensive margin condition which expresses that the net marginal benefit of residing in any city has to be equal across all cities which are inhabited. Condition (18) states that the utility level in the least developed city must be the same as the utility of living and working in the rural area.

Since (17) must also hold for the least developed city with productive amenity $\underline{A}_{it}^{p\,opt}$ we can use (17) and (18) to solve for this city's population level and then the Lagrange parameter:

$$\underline{N}_{it}^{opt} = \left(\frac{(\sigma+\eta)(1+\gamma)}{(\gamma+\theta)} \frac{B(\psi)^{1+\sigma}}{\tau_t} \frac{\underline{A}_{it}^{p\,opt} A_i^c \Lambda_i{}^{\gamma}}{(1-\varphi_i)} \right)^{\frac{1}{(\gamma+\theta)-(\sigma+\eta)}} \tag{19}$$

[19] Henkel et al. (2021) focus fully on the German fiscal transfer system, in contrast. However, even their analysis of optimal transfers targets only a constrained social optimum since they take local taxes as exogenously given.

$$\mu = \frac{(\gamma+\theta)-(\sigma+\eta)}{(\sigma+\eta)(1+\gamma)} \frac{\tau_t\,(1-\varphi_i)}{A_i^c\,\Lambda_t{}^{\gamma}}\,\underline{N}_{it}^{opt\,\gamma+\theta} \tag{20}$$

Comparing (19) and (20) with (8) and (9) shows that the least site to be populated in the social optimum is developed at its locally efficient scale, $\underline{N}_{it}^{opt} = N_{it}^{*}\left(\underline{A}_{it}^{p\;opt}\right)$ and that $\mu = v_{it}^{*}\left(\underline{N}_{it}^{opt}\right)$. Given the Lagrange parameter (20), the rural population follows from (16), $N_{rt}^{opt} = \left(\frac{A_{rt}}{\mu-\kappa_r}\right)^{\frac{1}{\lambda}}$, and the total city population is $N_t - N_{rt}^{opt}$. Socially optimal city sizes N_{it}^{opt} are implicitly determined by (17) after using μ from (20). Since (16) - (19) are analogous to the system in Albouy et al. (2019, Proposition 1), their proof can be invoked to show that socially optimal city sizes (for populated sites) have larger population than cities under local governments, and that fewer cities are inhabited in the social optimum. Note that the social optimum exhibits 'implicit land-use regulations': the equation of net marginal benefits across cities in the social optimum implies that cities with successively lower hump-shaped real income curves exhibit not only successively smaller population levels, but also successively lower different real incomes, and hence, successively smaller implicit 'land-use regulations' which only vanish for the last city.

Pulling the results of this section together we can state that the only (net) distortion arising under local governments and our chosen micro-foundations concerns city sizes which are too small and hence, cities are too numerous, relative to the social optimum. The intuitive reason is that local governments ignore the extensive margin of development which leads to the mentioned inefficiency when sites are heterogeneous: diminishing returns from lower site qualities at the system level which the social planner takes into account in the choice of city size are ignored by local governments (Albouy et al. 2019). Any potential distortion in city production vanishes on net due to the chosen micro-foundations of agglomeration economies (cf. footnote 5) and the choice of investment in human capital is undistorted as in Jones (2005).

3 Quantification

3.1 The data

The quantification of the model for the German economy necessitates that we estimate two pairs of key parameters for urban benefits (σ and η) and urban costs (γ and θ). To do so we follow DP (2022) and draw on structural relationships implied by the theoretical model. We start by characterizing our delineation of spatial units and by providing an overview over the three sets of microdata which inform us about individual earnings, rent prices and commuting behavior in German cities as well as complementary data-sources. Detailed characterizations of these data and definitions are relegated to appendix A. We focus on the time span between 1995 and 2017, i.e., we restrict our attention to developments after the German unification until the most recent year (2017) for which we can connect the various data to maintain consistency across our estimations.

BBSR-urban district regions. To accord with our labor market data which we take from the SIAB-R (Sample of Integrated Labour Market Biographies Regional File, see Antoni et al. 2019), our definition of cities draws on the well-established classification of district regions by the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR). These cities have a population greater than 100,000 inhabitants. Of these 328 district regions in Germany the BBSR specifies 264 as being located in 'urban areas' and we take these as corresponding to the cities in the model. The remaining population is assumed to locate in the ‚rural area'. Our population data for 2017 are taken from 'Regionaldatenbank Deutschland' provided by the Federal Statistical Office for Germany, the population data for 1995 come from ‚Laufende Raumbeobachtung des BBSR' provided by the BBSR. The resulting population dataset provides a time consistent definition for the territorial boundaries of the German districts - defined as on 31 December 2017 – so that the population of district regions in 1995 and 2017 can directly be compared (appendix A.1 provides details).

SIAB-R. Our estimates of the benefits of urban agglomeration - the static and the dynamic wage premium captured by the parameters σ and η -, draw on the mentioned SIAB-R, a panel dataset containing workers' employment histories across locations which allows us to control for unobservable individual heterogeneity in human capital. These data provide a 2% random sample of administrative social security records which comprehends employees subject to social security contributions and marginal part-time employment (Antoni et al., 2019:6). Since wages above the upper earnings limit for statutory pension insurance are top-censored and set equal to the upper earnings limit (Antoni et al., 2019:42), we exclude these observations from our sample. We also constrain our sample to full-time workers liable to social security with their workplace located in urban areas. Moreover, only individuals of German nationality that are at least 18 years old are included. This yields a final sample with 1,529,134 observations (see appendix A.2 for further details concerning the data preparation).

RWI-GEO-RED. To estimate the parameter γ of urban costs, we draw on the real-estate RWI-GEO-RED data provided by the Research Data Centre Ruhr at RWI (FDZ Ruhr; RWI, 2019). The dataset contains real estate advertisements with information on prices and additional characteristics of apartments and houses for sale and for rent obtained from the internet platform ImmobilienScout24 between 2007 and 2020. The location of an apartment is observed at a grid of 1-square-kilometer cells following the European standard ETRS89-LAEA (Schaffner 2019:6). To estimate the urban cost parameter γ we consider only apartments for rent and we focus on the year 2017 for the mentioned reasons. We consider only apartments located in urban areas and we restrict the sample to apartments with a living area between 50 and 400 square meters and rents between 50€ and 5,000€ (i.e., we exclude unusually cheap and luxurious objects). With information about the living areas and rent prices we obtain the rent per square meter which we use in our regressions (further details on the data and the data preparation are provided in appendix A.3).

Mobility in Germany. Our estimate of the congestion parameter in urban costs, θ, draws on the dataset 'Mobilität in Deutschland (MiD)' instructed by the Federal Ministry of Transport and Digital Infrastructure in Germany (BMVI) which provides information on the travel behavior of individuals in cities with residences observed at a very small spatial scale (BMVI 2017). This is a nationwide survey of households on their sociodemographic background and on their everyday traffic behavior in 2017. Randomly selected households are asked for their traffic behavior and distances travelled on a given reference date. The place of residence of an individual is observed on a grid of 1-square-kilometre cells that are assigned to the district regions in Germany. We prepare our dataset as follows. In order to capture the congestion arising from traffic, we restrict the sample to commuting trips by private car. Further, an interviewed person is only included if the person drove herself and is in her usual environment on the reference date such that the commuting tip is assigned to the correct district region of the individual's place of residence. Finally, the sample is restricted to drivers whose place of residence is in an urban area. This yields a sample of 57,034 individuals from 48,161 households with a total of 187,435 commuting trips. Further details on the dataset and the preparation steps are provided in appendix A.4.

Geographically constrained land. To quantify the share of developable land Λ_i, we follow Saiz (2010) and calculate for each city the share of geographically unconstrained land within the city's 30-km radius of the city center. Hereby, an area is classified as geographically constrained if it is covered by water and wetlands, slopes steeper than 15%, nature reserves or if it belongs to a foreign country (see appendix A.5 for further details). The spatial distribution of the geographical constraints is illustrated by fig. 2.

3.2 Population elasticity of urban benefits

Our strategy for the estimation of the two parameters of the benefits of agglomeration, the static and the dynamic earnings premium from working in a bigger city, σ and η, that are contained in the city wage equation (1) draws on the two-step methodology performed by De la Roca and Puga (2017). The estimation is enabled by the longitudinal worker-level information provided in the SIAB-R panel database. The regression framework consists of three equations:

$$ln\, y_{it}^{j} = a_i + a_j + a_t + \sum_i \; b_i\, e_{it}^{j} + X_t^j b + \epsilon_{it}^{j} \tag{21}$$

$$\hat{a}_i = \sigma \; \ln N_i + \epsilon_i \tag{22}$$

$$\hat{a}_i + \hat{b}_i \bar{e} = (\sigma + \eta) \; \ln N_i + \epsilon_i. \tag{23}$$

Figure 2: Geographical constraints

Notes: The map visualizes the spatial distribution of geographical constraints in Germany. Darker areas indicate district regions with a higher share of geographically constrained land, brighter areas indicate less geographically constrained regions.

Eq. (21) is a worker-level earnings regression which features log earnings of worker j in city i at time t as dependent variable. The independent variables consist of a city fixed effect a_i, a worker fixed effect a_j, a time fixed effect a_t, the experience e_{it}^{j} acquired by worker j in city i up until time t, a vector of time-varying individual and job characteristics X_t^j, and an error term ϵ_{it}^{j}. The scalars b_i and the vector b are parameters. The worker fixed effect in the regression controls for unobservable time-invariant characteristics of workers and the time fixed effect accounts for common (macroeconomic) factors that affect wages over time, e.g., inflation and technological progress. This first-step regression allows for a static earnings premium if the city fixed effect a_i is positively correlated with city size. This is analyzed with the second-step eq. (22) which regresses the estimated city fixed effects ($\hat{a}_i$) from the first-step regression on city size N_i, to obtain an estimate of σ. The first-step regression also allows for learning effects from working in big cities. An estimate of the dynamic earnings premium associated with learning (together with the static premium) is obtained from eq. (23). This is a further second-step regression, which regresses the estimated value of the experience accumulated in (bigger) cities $\hat{b}_i$ evaluated at the average local experience $\bar{e}$ along with the estimated city fixed effects from the first-step regression on city size N_i to obtain an estimate of $\sigma + \eta$. With the definition $\eta = \beta(1 + \sigma)$, the learning effect in cities β is then backed-out from the results of the two regressions (22) and (23).

Three remarks are in order before we proceed to the regressions and their results. First, it should be noted that the city fixed effects, which are crucial in this strategy, are only estimated from workers who change location between two dates, i.e., from 'movers'. If all workers stayed at the same workplace, it would not be possible to separately identify city fixed effects from worker fixed effects. This raises the issue that the estimate may be based on a possibly highly selected set of individuals (see Combes et al. 2011). Second, in taking the model to the data we face the decision of whether to base the estimates on city population or on population density (see Duranton and Puga 2020 for a discussion). Practical arguments associated with German district regions on which our city classification is based speak in favor of the use of city population data which we therefore use in the body of the paper (as explained in appendix B). However, in the appendix we also provide a robustness check based on population densities, which indicates that the results are very much in line with our specification in the body of the paper. Third, the experience variable needs to be operationalized, leaving much room for choice. We have opted to follow De la Roca and Puga (2017) in focusing on the experience in the 5 largest cities in terms of population in Germany which are Berlin, Hamburg, Munich, the Region Hannover and Cologne (and we focus on the year 2017 which yields $\bar{e} = 6.3$ years) in our data. We also consider the experience accumulated in German cities with a population of more than 500,000 excluding the mentioned five biggest cities in our regression.

Our regression results are shown in table 1. Column (1) presents the results of a one-step estimation of the parameter σ to which we turn later. The results of our focused two-step estimation are shown in columns (2) through (4). Column (2) provides the results of the first step regression eq. (21). It is seen that experience is more valuable when it is accumulated in bigger cities. The estimated coefficients for experience in the five biggest cities and in cities with a population exceeding 500,000 without the five biggest cities are both positive and significant, but the former (0.014) is stronger than the latter (0.011). Columns (3) and (4) of table 1 display the results of the second step regression eqs. (22) and (23). Column (3) shows an estimated elasticity of the static earnings premium with respect to city size of 0.017. This suggests that moving to a city of double the initial size increases earnings by 1.7%. Column (4) shows an estimated elasticity of the medium-term earnings premium with respect to city size of 0.037. Two-fifths of the medium-term premium are due to static benefits and three-fifths due to the dynamic benefits of agglomeration. With an estimated value for the parameter σ of 0.017, parameter η is estimated at 0.020. This yields an estimated value for the learning effect in cities β of also 0.020.

Table 1: Estimation of population elasticities of urban benefits

	(1)	(2)	(3)	(4)
Dependent variable:	Ln earnings		Static premium (city indicators column (2))	Medium-term premium (static + 6.3 years local experience)
Ln city size	0.01672*** (0.0011)		0.0172** (0.0079)	0.0373*** (0.0087)
City indicators		Yes		
Worker fixed effects	Yes	Yes		
Experience in five biggest cities	0.0139*** (0.0008)	0.0140*** (0.0008)		
Experience in five biggest cities x exp.	-0.0006*** (0.0001)	-0.0006*** (0.0001)		
Experience in cities > 500,000 (without five biggest)	0.0107*** (0.0007)	0.0100*** (0.0070)		
Experience in cities > 500,000 (without five biggest) x exp.	-0.0004*** (0.0001)	-0.0004*** (0.0001)		
Experience	0.0468*** (0.0007)	0.0462*** (0.0007)		
Experience^2	-0.0016*** (0.0001)	-0.0016*** (0.0001)		
Observations	1,529,134	1,529,134	264	264
R^2	0.3282	0.3382	0.0142	0.0600

Notes: All regressions include a constant term. Columns (1) and (2) include firm tenure and its square, year indicators, 4 occupational skill indicators, 14 sector indicators, and 120 occupation indicators. Column (2) in addition includes 264 city indicators. Worker values of experience and tenure are calculated on the basis of actual days worked and expressed in years. Coefficients are reported with robust standard errors in parenthesis, which are clustered by worker in columns (1) and (2). ***, **, and * indicate significance at the 1, 5, and 10 percent levels. The R^2 reported in columns (1) and (2) is within workers.

Our estimates of the two agglomeration parameters for Germany are quantitatively in line with results from similar regressions in the literature (see the survey by Combes and Gobillon 2015). Germany is also addressed in recent work by Grujovic (2021) and Dauth et al. (2022) who also use SIAB-data but different city samples and/or geographical units and who find very similar static premia of 0.014 and 0.017, respectively. De la Roca and Puga's (2017:120) estimates for Spain are also in the same ballpark. They find a parameter for the static earnings premium of 0.022 and for the dynamic premium of 0.029. In their study for the United States DP (2022:23) obtain significantly higher estimates of 0.045 for σ and of 0.032 for β.

Column (1) of table 1 presents the results of a one-step estimation of the parameter σ which draws on regression eq. (21) but replaces the city fixed effect with (the log of) population size, $\ln N_i$. This regression gives an alternative estimate for σ of 0.017 which is quite similar to the estimate that is obtained from the two-step procedure. As the latter addresses endogeneity issues more properly, we work with the results from columns (3) and (4).

3.3 Population elasticity of urban costs

Urban costs are represented by the two elasticity parameters of eq. (5) of the model, the elasticity of commuting costs (distance travelled) with respect to distance from the CBD, γ, and the elasticity of congestion costs/travel speed with respect to the city population, θ. We now address the estimation of these two parameters in turn.

Elasticity of commuting costs (distance travelled). Our strategy to estimate the elasticity of commuting costs (distance travelled) with respect to the distance from CBD follows DP (2022) and draws on eq. (4) which, along the lines of Alonso-Muth, implies:

$$\frac{d \ln \,[P_{it}(x) - P_{it}(0)]}{d \ln x} = -\frac{d \ln T_{it}(x)}{d \ln x} = -\gamma. \tag{24}$$

Eq. (24) shows that moving away from CBD, (quality-transfer) unadjusted housing costs must fall in the same proportion as commuting costs rise, and that this proportionate change is given by the parameter γ. Hence, with data on rents across locations within a city this parameter can be estimated by the following regression

$$\ln P_{il}^{j} = a_i - \gamma \ln x_{il}^{j} + X^{j} b + \epsilon_i^{j} \tag{25}$$

where $\ln P_{il}^{j}$ is the log rent/m² for renter-occupied apartment l in grid cell (block) j, $\ln x_i^j$ is the log of distance between the place of residence and the city-center, a_i is a city fixed effect and

X^j is a vector of dwelling and neighborhood characteristics, b is a parameter vector and ϵ_i^j is an error term.

Table 2: Estimation of elasticities of urban costs

	(1)	(2)
Dependent variable:	Ln rent price per square meter	Ln estimated city travel speed
Ln distance to city center	-0.0709***	
	(0.0003)	
Ln city size		-0.0680***
		(0.0062)
City indicators	Yes	
Controls	Grid cell & dwelling characteristics	
Observations	484,812	264
R^2	0.7404	0.2794

Notes: All regressions include a constant term. Column (1) includes city indicators and, as grid cell controls, indicators for riverfront and oceanfront location, and average unemployment rate and purchasing power per household. Average unemployment rate and purchasing power are centered at the city mean. Dwelling controls comprise living area, indicators for the number of rooms, object category, and construction decade. In column (2), the dependent variable is the log of city travel speed estimated in a previous step by regressing travel speed for individual trips by private car on city indicators, including the same grid cell controls as in column (1) in addition to driver and trip controls. We use this to predict for each city the speed of a 15km commuting trip on a Tuesday at 8AM by a driver with average characteristics. Coefficients are reported with robust standard errors in parenthesis. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

We use the RWI-GEO-RED data (RWI 2019; Schaffner 2020) and proceed as follows. The distance of an apartment to the center of its district region is calculated as the haversine distance between the midpoint of the grid cell in which an apartment is located and the center of the district region which we define as the place indicated by Google Maps for the core of the main city of the district region. The dwelling characteristics that we include as controls comprise the living area, indicators for the number of rooms, object category, and the construction decade. Table A4 in appendix A.3 provides descriptive statistics on these dwelling characteristics as well as on the rent per square meter, the dependent variable in regression eq. (25). The neighborhood characteristics that we take into account are the average unemployment rate and the purchasing power per household in a grid cell[20] - both centered at the city mean -, as well

[20] The data for both come from RWI and microm (2017a) and (2017b) provided by the Research Data Centre Ruhr at RWI. The riverfront location is defined as the location of a grid cell at a large river, whereby large rivers have a catchment area larger than 50,000 km^2 or other rivers and tributaries that have a catchment area larger than 5,000 km^2. The corresponding shapefile is from the European Environmental Agency (Permalink: D7925F3C-AFF7-4256-8162-513A2C1C69E3), as is the shapefile for the construction of the indicator for oceanfront location (Permalink: 88055d120fd54c82a3606b97502d21c1).

as indicators for oceanfront and riverfront location. The city fixed effect a_i absorbs $\ln P_{it}(0)$, the natural log of the rent per square meter of a national-reference apartment at the center of its district region for city-average grid cell characteristics in eq. (25) whilst $\ln P_{il}^j$ corresponds to $\ln P_{it}(x)$. Column (1) of table 2 shows our regression results. As can be seen, we find an estimate of 0.071 for the parameter γ in Germany. Hence, rent prices per square meter in an urban district region decline by 7.1% when the distance to the city center of the region is doubled. Duranton and Puga (2019:25) obtain a similar estimate (0.0734) for the United States.[21]

Elasticity of congestion/travel speed. Our estimation strategy for the elasticity of congestion costs/travel speed with respect to the city population θ is also inspired by DP (2022). Recall that the cost per unit commuting distance in the model (which can be thought of as time loss in commuting) is given by $\tau_{it} = \tau_t\, {N_{it}}^{\theta}$, or, rewritten in logarithmic form, $ln\, \tau_{it} = ln\, \tau_t + \theta\, ln\, N_{it}$. To a first approximation τ_{it} can be taken to be inversely proportional to the travel speed $\hat{\tau}_i$ in the city, i.e., $\tau_{it} \sim 1/\hat{\tau}_i$. This then inspires the regression:

$$\ln \hat{\tau}_i = b + \theta \ln N_i + \varepsilon_i \tag{26}$$

where b is a constant, ε_i is an error term, and θ is the key parameter which can be estimated with data on travel speed in cities (and city population). The survey of household traffic behavior, Mobility in Germany (BMVI 2019), allows us to perform this exercise.

We proceed in two steps. In the first step, we estimate the travel speed in each city specified as the speed of a 15km commuting trip on a Tuesday at 08:00AM by a driver with average characteristics,

$$\ln \tau_{ik}^j = a_i + X_{ik}^j b + \varepsilon_{ik}^j \tag{27}$$

where τ_{ik}^j is the travel speed of driver j for the commuting trip k in city i, a_i is a city fixed effect, X_{ik}^j is a vector of grid cell, driver and trip characteristics, b is a vector of parameters and ε_{ik}^j is an error term. We follow DP (2022:43) in using the following controls for grid cells: water and riverfront location, the unemployment rate, and the purchasing power of a household. Trip characteristics contain the natural logarithm of the trip distance, indicators for the day of week and departure time in 30-minute intervals, and the trip purpose. The driver characteristics

[21] Note that we purposefully report the estimate stated in the predecessor version of DP (2022) since our estimation strategy follows this earlier work. Using a related but slightly different approach DP (2022) find a value of 0.0769.

comprise indicators for age groups, single-person household, for retires, for the driver being male, the household structure, the number of drivers in the household, and the driver's distance to the city center (descriptive statistics on these control variables are included in table A5 in appendix A.4). Based on the estimated parameter vector $\hat{b}$ and the estimated city fixed effects $\hat{a}_i$, we calculate the speed of a 15km commuting trip on a Tuesday at 08:00AM by a driver with average characteristics for each city i, $\hat{\tau}_{ik}^j$.

In the second step, we use this estimate as dependent variable and regress it on a constant and the city population N_i as in eq. (26) where we use $\hat{\tau}_{ik}^j$ for $\hat{\tau}_i$. Our results are shown in column (2) of table 2. We obtain a highly significant estimate of 0.068 for the parameter θ (a standard error of 0.0062) and a regression R-squared of 0.279 for Germany. This implies that with a doubling of the city population the travel speed in the city falls by 6.8%. Our results compare with an estimated value of the elasticity of travel speed of 0.0388 for the United States (DP 2022:22).

3.4 Population elasticity of rural income

A further important parameter is the population elasticity of rural income which plays a key role for the rural wage in eq. (2). Since we lack data for the estimation of this parameter, we take it from the study of factor shares for the United States by Valentinyi and Herrendorf (2008). They find a value of $\lambda = 0.18$, which we adopt for our model calibration and counterfactuals. To check robustness, we have scaled this parameter up and down around this value. None of our results is crucially affected by these variations.

4 The German City System and the Aggregate Economy

4.1 The German city system through the lens of the model

Calibration. We now analyze the German city system through the lens of the model focusing on the situation in 2017. We assume in the following that the factual Germany city sizes in this year are indeed a general equilibrium of the model under local governments, i.e., factual city sizes in 2017 correspond to N_{it}^* as specified by equation (8). In the spirit of DP (2022) we assume that the cities' incumbents are represented by the city populations in the initial year of our analysis (the year 1995) and we take the citizens entering cities from 1995 to 2017 to be the newcomers (if positive). We use the following approximations for our estimated parameters, $\sigma = 0.017$, $\eta = 0.020$, $\gamma = 0.071$ and $\theta = 0.068$, and we assume $\lambda = 0.18$.

The model postulates that cities are populated in descending order of real incomes v_{it}^*. To quantify real incomes, we start by backing out productive amenities A_{it}^p from the wage equation (1) using the observed wages and city sizes. We scale the value of fiscal transfers across cities from the various German fiscal equalization schemes (Henkel et al. 2021) by net urban costs (5) and observed city sizes to obtain the fiscal transfer rates φ_i.[22] With this information at hand, we filter out consumption amenities A_i^c from the equation for optimal city sizes (8).[23] This allows us to calculate v_{it}^* from (9).[24] The productivity parameter in the rural area is calculated from the condition that the real income in the marginal city has to be equal to the real income in the rural area, $\underline{v}_{it}^* = v_{rt}^*$, which can be rearranged to yield $A_{rt} = \left(\underline{v}_{it}^* - \kappa_r\right) N_{rt}^{*\,\lambda}$.

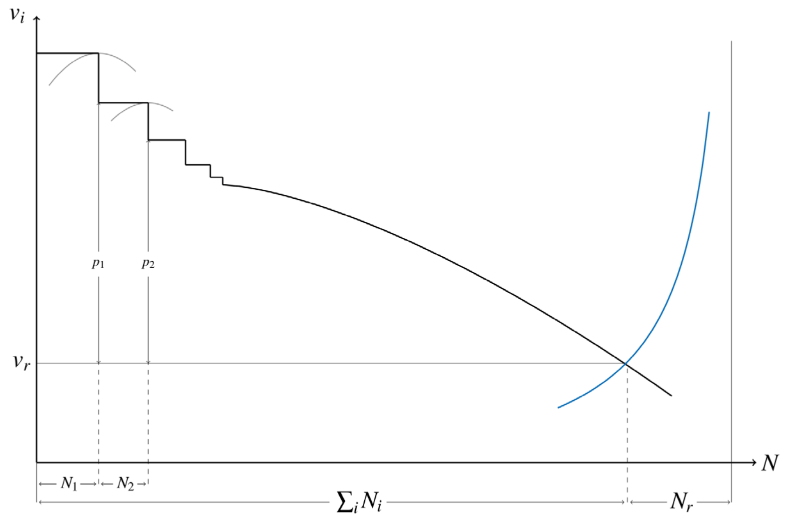

Figure 3: The German city system in 2017

Notes: The figure illustrates the allocation of population across German urban and rural areas in 2017 as the equilibrium of the model under local governments. The vertical axis depicts real incomes, and the length of the horizontal axis is total population in Germany with part of the population living in cities $\sum_i N_i$ and part living in the rural area N_r. Real income in the rural area as a function of the rural population is represented by the blue curve depicted from the right to the left. The thick horizontal segments give the equilibrium real income for incumbents in each city and the length of the segments is the population of the corresponding city. Equilibrium real incomes result from the maximization of real income with respect to city population by local governments represented by the thin gray curves. German cities are plotted in descending order of real incomes. The spatial equilibrium of the model is determined by the intersection of the black downward-sloping curve (locally optimal real income peaks) with the blue curve (real income in the rural area).

[22] Henkel et al. (2019) describe the institutional background of the German fiscal transfer system, a complicated set of rules comprising various fiscal equalization schemes, and they meticulously work out local tax revenues before and after redistribution, and so the net transfers across locations. We are very grateful to these authors for generously providing us these data which allowed us to calculate the fiscal transfer rates for our analysis.

[23] Appendix C.1 provides details and results for productive and consumptive amenities and the fiscal transfer rates.

[24] Our quantification yields real income levels relative to the normalized commuting cost factor τ_t which prevails in all cities and which we assume to be 10,000.

Fig. 3 portrays the German city system in 2017 schematically. The total population of 82,52 million in 2017 is depicted on the horizontal axis.[25] German cities (urban district regions) are listed in descending order of real incomes. The city with the highest real income level, Ingolstadt, is depicted as city 1 with $N_1 = 135{,}244$. Boeblingen comes second with a population of $N_2 = 389{,}548$, and so on for all 264 urban districts. Munich comes at position 7, Hamburg at position 35, and Berlin at position 139. The total urban population, $\sum_i N_i$, can be read off the intersection of the black downward-sloping curve connecting locally optimal real income peaks with the blue curve which shows the declining marginal product of labor in the rural area (plus the fiscal balance term) from right to left. The rural population amounts to $N_r =$ 9 million. Also shown are the land-use regulations p_{it}^* that are assumed to be levied by city incumbents on city newcomers. The highest level of these land-use regulations p_{1t}^* is imposed by the most attractive German city, Ingolstadt, followed by Boeblingen, whose citizens impose land-use regulation amounting to p_{2t}^* in terms of output. The marginal city, Erzgebirgskreis, imposes no land-use regulations, $p_{264t}^* = 0$.

Empirical support. Several independent pieces of evidence provide strong empirical support for our theoretical model and calibration. Our model innovations, notably the incorporation of consumption amenities and transfers, are key for the capacity of the model and calibration to accord with three of the four sets of facts that prevail in Germany as we now show.[26]

First, the DP (2022) model and our modification imply that average human capital per worker is greater in bigger cities (cf. section 2.1, human capital formation), but that the growth rates of human capital, a key component of city growth, do not systematically vary across cities of different size (cf. eq. 12). This implication is borne out by the facts for Germany as shown in fig. 4, panel A. The evolution of the share of workers with a college degree for three size classes of urban district regions is characterized by different levels but parallel trends.

A crucial second piece of evidence externally validates the model-implied real income (indirect utility) ranking of Germany's cities depicted in fig. 3. Panel B of fig. 4 shows the very close correlation between our measure of real income based on the local government allocation with our model and data (vertical axis) and an independent measure of real income (horizontal axis) which uses the regional price indices recently established by Weinand and Von Auer (2020) on

[25] Germany has experienced only small population growth in the time span considered, so the general equilibrium for 1995 (where German population was at 81,45 million) looks much the same.

[26] Further model implications could be confronted with the data, if systematic information on land-use regulations, similar to the Wharton Residential Land Use Regulatory Index (Gyourko, Saiz and Summers 2008; Gyourko, Hartley and Krimmel 2021) was available for Germany, which unfortunately is not the case.

the basis of micro-price consumer data collected for German counties by the Federal Statistical Office and the Statistical Offices of the Länder as deflators for the average wages in cities. The correlation coefficient between the two measures exceeds 0.9. The positioning of individual cities in the implied city ranking is also very similar, as illustrated by Berlin which comes at position 139 in our ranking and at position 176 in the alternative ranking.[27]

Third, the political-economy mechanism of city-size determination developed by DP (2022) comes with the novel implication that, due to the land-use regulations imposed by incumbents on city newcomers, house prices at the edges of cities are not equated, in contrast to the standard monocentric city model.[28] Rather, there are discrete differences, and these are systematically higher, the more attractive the city (cf. eqs. 10 and 11).[29] The scatterplot of periphery rents for all urban districts in our database presented in panel C of fig. 4 shows that this implication accords well with the German facts.

Fourth, panel D in fig. 4 presents the evolution of housing rents over the extension of cities for a convenient selection of cities, Berlin, Hamburg, and Munich, to exemplify important further points. First, the gradient of housing rents is falling which accords with the predictions of the standard monocentric city model, the backbone of our urban model (cf. section 2.1). More importantly, it is also seen that the rents at the city edge in Munich, exceed those at the city edge of Hamburg, and that Berlin, the much-featured biggest German city which has yet a lower real income (attractivity) than the previous two, exhibits yet lower periphery rents, reiterating the point we made in our discussion of panel C.[30] Panel D shows another important finding. Our model predicts that the (quality-transfer) adjusted house prices in the city center are a sufficient statistic for the attractiveness of a city. These adjusted prices are the relevant model predictors for the empirically observed house prices at the city centers. The data shown in panel D accord with this implication. Munich exhibits the highest factual rents in the city center, followed by Hamburg. Berlin, the other city in our selection, has the lowest city center rents.

---

[27] To establish the figure shown in fig. 4 Panel B, we adjust the regional classification (from counties to district regions) as explained in Appendix A1. Note that the alternative measure does not take consumption amenities and fiscal transfers into account (which shifts the city Berlin up in our ranking, for example).

[28] DP (2022) provide an extended discussion of the differences and similarities of their model and the standard monocentric city model.

[29] Our modification of the DP (2022) model plays out here. Abstracting from land constraints there is a 1:1 relation between city size, city productivity and city attractiveness (real income), in DP (2022) so that they can largely focus on the relationship between periphery prices and city size. In our modification of the model, consumptive amenities and fiscal transfers are further determinants of a city's attractiveness which explains why we focus on attractivity (not city population) in figure 4, panel B (and also in panel C).

[30] It should be noted that, due to the different components of a city's attractiveness in our modification of the model alluded to in the previous footnote, there is no 1:1 relationship between the ranking of housing prices and the extension of cities which is determined by the city population, in contrast to DP (2022).

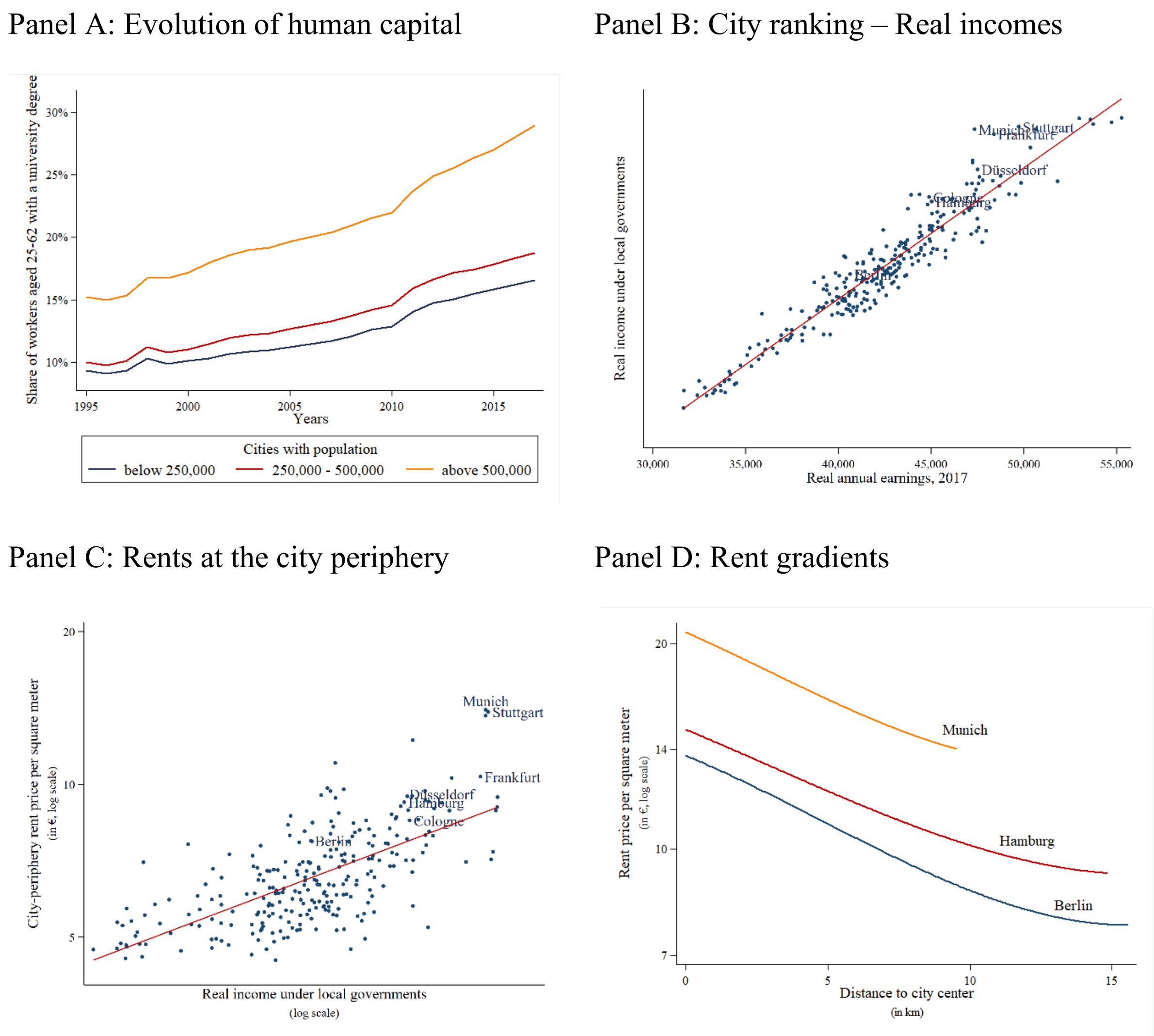

Figure 4: Empirical evidence

Notes: Panel A shows the evolution of the share of workers aged between 25 and 62 with a university degree using the SIAB regional file. Each city is assigned to one of the three population size categories. Panel B plots real incomes under the local government allocation of our model and data (vertical axis) against an independent measure of real income (horizontal axis) which uses the regional price indices recently established by Weinand and Von Auer (2020) on the basis of micro-price consumer data collected for German counties by the Federal Statistical Office and the Statistical Offices of the Länder as deflators for the average wages in the cities. The bridge from counties to district regions is explained in Appendix A1. Panel C plots city-periphery rent prices per sqm in 2017 against real income under local governments. The city periphery is defined as the 95th percentile of dwelling distances from the CBD. The rents at the city periphery are estimated from a regression of the log of the rent per sqm on a third-degree polynomial of distance to the CBD allowing the coefficients of the polynomial to be city specific. The regression includes the same controls as column (1) of table 2. Panel D depicts the rent price per sqm as a function of distance to the CBD for the three cities Berlin, Hamburg, and Munich. The rent gradients are obtained from the regression used to establish panel C. City-periphery rents in panel B and rent gradients in panel D are plotted for a dwelling with average national characteristics in a neighborhood with average city characteristics using the RWI-GEO-RED data.

4.2 Cities and aggregate income in Germany

Expansion of the Top Seven. Our theoretical analysis of the welfare properties of the model shows that existing cities are undersized under local governments (cf. section 2.3). This conforms with the public perception that housing in Germany's most attractive cities, notably the Top Seven Berlin, Hamburg, Munich, Cologne, Frankfurt, Stuttgart, and Duesseldorf is dramatically short of demand (cf. section 1). We therefore now consider a counterfactual expansion of the city populations in these cities. We assume a proportionate increase of each city by 10%, so that $\widehat{N}_{it} = 1.1\, N_{it}$, where we us a hat ($\widehat{\ }$) to indicate the counterfactual city population (and the counterfactual value of any other variable in the following). Given the factual (baseline) population in these cities in 2017, this implies that Berlin expands by 360,000 citizens, Hamburg by 180,000, Munich by 150,000, Cologne by 110,000, Frankfurt by 75,000, Stuttgart by 63,000 and Düsseldorf by 62,000, so that a total of 1 million people become new citizens in these locations. This counterfactual command that land-use regulations are reduced. Given the long-run nature of the model this can be understood as an increase of the housing supply in the Top Seven.[31] We assume that all migrants to these cities become newcomers in their destinations. Remember at the outset of the following analysis that, in the initial equilibrium, the real incomes of city incumbents v_{it}^* exceed the real income of city newcomers $v_{it}^* - p_{it}^* = v_{rt}^*$, which coincide with real incomes of rural workers $v_{rt}^* = A_{rt}\, N_{rt}^{*\,-\lambda} + \kappa_r$.

To analyze the consequences of this counterfactual we start by providing a qualitative description of the implied population reallocation. The algorithm that we use to implement this counterfactual is outlined in appendix C.2 Note that incumbents in cities other than the Top Seven have no incentive to change their city sizes because they remain locally optimal. Hence, incited by the lower land-use regulations, the first migrants to the Top Seven must come from the rural area. This fall in the rural population raises the marginal product of labor and so the real income in the rural area which now starts to exceed the (initial) real income of incumbents in the least productive cities. Consequently, the cities with the lowest real income will successively be vacated and their former residents migrate along with rural workers to Berlin, Hamburg, Munich, Cologne, Frankfurt, Stuttgart and Duesseldorf. This process continues until the population of the Top Seven has increased by 10% and the spatial equilibrium condition is restored. In this new equilibrium the real income of rural workers and citizens in the new marginal city are again equalized. Note that incumbent residents in cities other than the Top Seven, even though they maintain their city sizes, will adjust (lower) their land-use regulations to correspond with the new (and higher) threshold of real income. The described reallocations

[31] This counterfactual is, grosso modo, in the ballpark of the 'housing shortage' that is currently diagnosed in a recent study commanded by various German Associations, including the Labor Union IGBAU, the Tenants' Association (Deutscher Mieterbund) and Caritas (Pestel 2023). To put our counterfactual into the perspective it is also instructive to note that the current German government plans to expand the housing supply by 400,000 housing units per year for "some years", but the current record falls dramatically short of this, FAZnet, January 12, 2023 (https://www.faz.net/aktuell/politik/studie-zum-wohnungsmangel-700-000-wohnungen-fehlen-18598219.html).

and the new equilibrium can be imagined by reference to fig. 3. The counterfactual shifts the black downward-sloping 'ladder'-curve right- and upwards, so that at the new intersection, i.e. equilibrium, the total urban population increases, the rural population shrinks and the now higher real income in the rural area, $\hat{v}_{rt}$, defines the new utility threshold for the development of the (new) marginal city which is also the new real income of city newcomers. Our calculations show that the population in the rural area falls from initially 9 million to 8.3 million and that one city (Erzgebirgskreis) would be vacated, so that 263 cities remain.

The implied consequences for real incomes vary across groups. Incumbents in the Top Seven see their real incomes fall because their cities are now expanded beyond the locally optimal sizes. We calculate that their real income falls by 0.002%. The intuition for this mild real income loss is that the ∩-shaped real income curve is flat near the optimum, where agglomeration benefits and costs are balanced from the perspective of local governments. Incumbents in the 256 remaining cities other than the Top Seven maintain their real incomes, as already mentioned. All other groups gain in the counterfactual scenario, and they do so in a substantial way. Residents remaining in the rural area gain because of the increase in the marginal product of labor and so do newcomers in all cities because of the implied reduction in land-use regulations. This common real income gain amounts to 1.26%. The real income gain for former incumbents moving from the vacated city to the Top Seven also amounts to 1.26%. Note that these gains are constrained by the fact that land-use regulations in the new equilibrium are still substantial, which itself is a consequence of the mentioned flatness of real income curves near the optimum: incumbents in the Top Seven have to adjust the land-use regulations only by a small amount to accommodate their city expansions. Weighting all gains and losses with the respective population share we derive the aggregate consequences for real income in Germany for this counterfactual: average real income per person increases by 1.12%. The analysis thus substantiates the notion that the urban containment policies of incumbents in the Top Seven in Germany have significant societal costs.

To put our results into perspective we note that large population movements and welfare effects in the same ballpark have also been found in related research addressing the US city system exploring similar counterfactuals with partly quite different models.[32] We are not aware of another study performing a similar analysis for Germany, however.

[32] See DP (2022), Desmet and Rossi-Hansberg (2013) and Davis et al. (2014). Substantially larger welfare effects associated with the reduction of various land-use regulations are reported by Turner et al. (2014) and, notably, by Hsieh and Moretti (2019), surprising findings as noted by Proost and Thisse (2019), that merit further research.

Counterfactual shift to the social optimum. We have established the social optimum in section 2.3 to assure ourselves of the distortions and the directions of these distortions under local governments, of which there are potentially several in the model. As shown and discussed, the only (net) distortion arising under our chosen micro-foundations of agglomeration economies concerns city sizes which are too small under local governments, since they ignore the extensive margin of development, which leads to the mentioned inefficiency when sites are heterogeneous. This incited us to explore the expansion of a set of most attractive cities. One may also wonder about the effects of a complete switch from the status quo local government allocation to the social optimum. Such a counterfactual would be extreme, however, commanding a dramatic lowering of land-use regulations of 95.7% on average, and it would imply a hefty fall of the rural population down to 346,00 people and the vacation of 253 cities, as our calculations show.[33] These findings do not come as a surprise, however, as similarly strong results are also reported by DP (2022) and the preceding version Duranton and Puga (2019), and even more dramatic ones by Hsieh and Moretti (2019), for counterfactuals that are (partly) much less comprehensive than a shift to the social optimum. We take these extreme findings as suggestive of the need to explore the makings of the land-use regulations and the role of idiosyncratic preferences as mobility brakes more closely in future work.

Removing fiscal transfers. As a side-effect of our incorporation of fiscal transfers into the model we can also address their locational and welfare impacts. This allows us to gain a perspective on results that have been obtained in extant research which addresses the German fiscal transfer system head on with an entirely different model and methodology (Henkel et al. 2019). We therefore consider now the complete abolishment of fiscal transfers across cities in another counterfactual. To obtain the locally optimal city sizes, the corresponding real incomes, and the land-use regulations without fiscal transfers, we set $\varphi_i = 0$ and $\kappa_r = 0$ in eqs. (8) – (11). Cities that received fiscal transfers loose population whilst the citizenship in previous donor cities rises and residents shift from the former to the latter. It turns out that the total urban population rises when fiscal transfers are abolished. Hence, part of the former rural population along with former residents from cities with the lowest real incomes, that will successively be vacated, migrate to cities with higher real incomes. In the new (counterfactual) equilibrium, the rural real income is again equal to the real income in the marginal city and the spatial

[33] This extreme scenario would also imply a substantial rise of average real income per person by 34.44%. Incumbents in the remaining 11 cities experience real income losses per person between 1.01% and 4.18% which are mild relative to the big real income gains of newcomers and the rural population (+72.54%), and incumbents of vacated cities (+0.35% to 72.54%). Appendix C.3 provides further details.

equilibrium is retained by the corresponding adjustments of land-use regulations. The population in the rural area falls from initially 9 million to 3.7 million and 17 cities would be vacated, so that 247 cities remain. We assume that all residents that migrate to a new city become newcomers at their new destinations. It is straightforward to derive the real income changes for the various groups of citizens, see appendix C.4 for details. We focus here on the overall effect which we obtain by aggregating the weighted real income change of each group using the respective population shares as weights. Real income per person increases by 2.60% when fiscal transfers are abolished. By implication, status quo fiscal transfers in Germany have an overall negative welfare effect. This can be rationalized as follows. Our analysis of the social optimum indicates that cities should be expanded relative to the local government allocation and that fewer cities should be inhabited. The counterfactual quantitative analysis shows that the removal of fiscal transfers involves such a population shift away from the rural area and cities with the lowest real incomes towards cities with higher real incomes, the former donors, and thereby shifts the city system in the direction of the social optimum. Our analysis thus leads to an interesting difference to the study of Henkel et al. (2019). They find that the German fiscal transfer system raises aggregate welfare. To understand why this is the case it is important to note that they conduct their analysis not only with an entirely different model, a quantitative spatial model with a fixed number of cities in the spirit of Allen and Arkolakis (2014) and Fajgelbaum et al. (2019), which does not allow for the fundamental trade-off between agglomeration and congestion forces and which also does not allow for an extensive margin of city development. The further and crucial difference is that Henkel et al (2019) assume an entirely different population allocation mechanism, a free migration equilibrium (without land-use regulations), in contrast to the local government allocation stipulated in our analysis and backed up by the empirical evidence discussed in section 4.1.[34]

4.3 The Contribution of Cities and Agglomeration Economies to Growth in Germany

We now turn to the implications of urban growth for aggregate growth in Germany. Taking the log-difference of the expected value of income per person across two points in time, one obtains from eq. (1):

$$\mathbb{E}(\Delta \ln w_{it}) = \mathbb{E}\big(\Delta \ln A_{it}^{p}\big) + (1+\sigma)\,\Delta \ln B(\psi_t) + (\sigma+\eta)\,\mathbb{E}(\Delta \ln N_{it}) \tag{28}$$

[34] In the free migration equilibrium without fiscal transfers considered by Henkel et al. (2019), the 'core' (i.e. biggest) German cities are too big and the fiscal transfer system alleviates this inefficiency by shifting economy activity away from core cities towards peripheral areas.

The first component on the right-hand side of eq. (28) captures the growth of productive amenities. Crucial for our analysis are the second and third term. The static agglomeration parameter σ magnifies the impact of human capital accumulation on per person income growth from 1 to $(1+\sigma)$. Moreover, with $\sigma+\eta>0$, city population growth directly contributes to income growth. Expected city growth can similarly be derived from log-differencing eq. (8):

$$\mathbb{E}(\Delta \ln N_{it}) = \frac{1}{(\gamma+\theta)-(\sigma+\eta)}\left[\mathbb{E}\left(\Delta \ln A_{it}^{p}\right) + (1+\sigma)\Delta \ln B(\psi_t) - \Delta \ln \tau_t\right] \qquad (29)$$

This shows that expected city growth is affected by the development of travel costs, human capital accumulation, and productive amenities, and crucially so by the agglomeration parameters $\sigma+\eta$ whose size is key for the 'multiplier' that scales the impact of these drivers.

Eq. (28) can be brought to the data to gauge the impact of cities and agglomeration economies on growth in Germany. For this, we need data on average income growth per capita, growth in human capital accumulation and average city growth. We have the following information. In the time span from 1995 to 2017, average growth in income per person is 1.4% (Destatis 2021:21), so $\mathbb{E}(\Delta \ln w_{it}) = \ln(1.014)$, average city population growth is 0.2% per year (see appendix A.1), so $\mathbb{E}(\Delta \ln N_{it}) = \ln(1.002)$, and human capital grew by 0.2 % per year (Penn World Tables PWT 10.0), hence $\Delta \ln B(\psi_t) = \ln(1.002)$.[35]

Our counterfactual is to decrease agglomeration economies until they are eliminated. Start with the role of human capital formation. With static agglomeration represented through $\sigma = 0.017$, $\Delta \ln B(\psi_t)$ is multiplied by 1.017, so that the contribution from human capital formation to the annual growth rate of income per person is magnified from 0.2 to 0.203 percentage points. Hence, static agglomeration benefits raise the impact of human capital accumulation on growth in income per person by 0.003 percentage points. With urban agglomeration economies acting through static and dynamic benefits $\sigma+\eta = 0.037$, city population growth contributes $(\sigma+\eta)\ \mathbb{E}(\Delta \ln N_{it}) = 0.007$ annual percentage points to per capita income growth per year.[36] The overall effect of agglomeration economies on per capita growth is obtained by adding up the 0.003 percentage points from the magnification of human capital formation and the 0.007 percentage points contributed by city growth to yield additional 0.011 percentage points of income growth per person.

[35] With these data, we can also back out the growth rate of productive amenities from eq. (28) as 1.2% per year, i.e. $\mathbb{E}\left(\Delta \ln A_{it}^{p}\right) = \ln(1.012)$. The growth rate of travel costs can be inferred from eq. (29) as 1.4% per year, i.e. $\Delta \ln \tau_t = \ln(1.014)$.

[36] We also note that as $\sigma+\eta$ approaches nil in eq. (29) city population growth is reduced from the observed 0.2% to 0.12% per year.

Hence, agglomeration economies and cities resulted in a 0.24% higher output in the 22 years from 1995 to 2017, one quarter of which is due to their impact on human capital accumulation and three quarters of which are due to city population growth acting through static and dynamic benefits. The growth effect of urban agglomeration economies for Germany is milder than what DP (2022) report for the United States. This is unsurprising because we find much smaller static and dynamic agglomeration economies for Germany – consistent with extant research cited above. These elasticities for Germany are less than half those reported by DP (2022).[37]

5 Conclusions

This paper extends the urban growth model of Duranton and Puga (2022) to explore the impact of cities on local firms and households and the aggregate economy of Germany.

We innovate on the model by adopting alternative micro-foundations for agglomeration economies and a non-linear specification of human capital accumulation. This allows us to characterize the social optimum and to bring the model in line with semi-endogenous growth. A further key theoretical innovation is our incorporation of consumptive amenities and fiscal transfers.

On the empirical side we exploit the structural equations of the model and rich sets of micro-data for Germany's labor markets, housing rents, and household travel-to-work data within cities to estimate the elasticities of urban benefits and costs. We are the first to establish elasticities for urban costs for Germany, an estimated elasticity of commuting costs with respect to distance travelled of 0.071, and an estimate for the population elasticity of travel congestion of 0.068. Our estimates for static and dynamic agglomeration elasticities are 0.017 and 0.020, respectively. All estimates are highly significant.

We also innovate on the calibration strategy to capture the important role of consumptive amenities and fiscal transfers in Germany. The model and our innovations and calibration receive strong support from several independent pieces of evidence, the evolution of human capital, rents at the city periphery, rent gradients over the extension of cities, and consumptive amenities.

Our key policy counterfactual is a proportionate increase of the population of 10% in Germany's Top Seven metropolises which implies a significant overall welfare benefit of 1.12% per person which involves mild losses for city incumbents but strong gains for city

[37] Our analysis neglects that cities also foster innovation which would raise agglomeration economies, the same is true for DP (2022:35), however, so the country difference must have other sources.

newcomers. This indicates that urban containment policies in the Top Seven in Germany have significant societal costs. We also address a counterfactual shift to the social optimum which would imply yet stronger population movements and welfare gains and a counterfactual removal of all fiscal transfers across cities which would come with an aggregate welfare gain. Our final exercise is an evaluation of the implications of cities and agglomeration economies for aggregate growth in Germany. We find that these account for 0.011 additional percentage points of growth in income per capita per year.

We see several avenues for future work. One general avenue concerns the inclusion of idiosyncratic location preferences into the theoretical framework, their empirical identification, and an empirical assessment of their effects. A second avenue concerns the integration of further more detailed aspects of the supply side of the housing markets, including the key land-use regulations. Another avenue is to extend the model and analysis to address the important concern about skill and qualification differentials in real incomes in and across cities.

References

Abdel-Rahman, Hesham M. and Masahisa Fujita (1990). Product variety, Marshallian externalities, and city sizes. Journal of Regional Science 30:2, 165-183.

Ahlfeldt, Gabriel M. and Elisabetta Pietrostefani (2019). The Economic Effects of Density: A Synthesis. Journal of Urban Economics 111, 93-107.

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf (2015). The Economics of Density: Evidence from the Berlin Wall. Econometrica 83:6, 2127-2189.

Albouy, David, Kristian Behrens, Frédéric Robert-Nicoud and Nathan Seegert (2019). The optimal distribution of population across cities. Journal of Urban Economics 110, 102-113.

Allen, Treb and Costas Arkolakis (2014). Trade and the Topography of the Spatial Economy. Quarterly Journal of Economics 129:3, 1085-1140.

Antoni, Manfred, Andreas Ganzer and Philipp vom Berge (2019). Sample of Integrated Labour Market Biographies Regional File (SIAB-R) 1975 - 2017, Institute for Employment Research, FDZ-Datenreport, 04/2019 (en).

Behrens, Kristian and Frédéric Robert-Nicoud (2015). Agglomeration theory with heterogeneous agents, in: Handbook of Regional and Urban Economics 5A, G. Duranton, J.V. Henderson and W. C. Strange (eds.), 171–245.

Bils, Mark and Peter Klenow (2000). Does Schooling Cause Growth? American Economic Review 90: 1160-1183.

Black, Duncan and J. Vernon Henderson (1999). A Theory of Urban Growth. Journal of Political Economy 107:2, 252–284.

BMVI (2017). Mobilität in Deutschland 2017, Bundesministerium für Verkehr und digitale Infrastruktur.

Brueckner, Jan K., Jacques-Francois Thisse, and Yves Zenou (1999). Why is central Paris so rich and downtown Detroit poor? An amenity-based theory. European Economic Review 43: 91-1007.

Bundesinstitut für Bau-, Stadt- und Raumforschung (2010). Gebietsreformen – politische Entscheidungen und Folgen für die Statistik. BBSR-Berichte KOMPAKT.

Bundesinstitut für Bau-, Stadt- und Raumforschung (2012). Raumabgrenzungen und Raumtypen des BBSR. Analysen Bau.Stadt.Raum Vol. 6.

Bundesinstitut für Bau-, Stadt- und Raumforschung (2016). Dokumentation Zensuskorrektur der Bevölkerungszeitreihe.

Combes, Pierre-Philippe, Gilles Duranton and Laurent Gobillon (2011). The Identification of Agglomeration Economics. Journal of Economic Geography 11, 253-266.

Combes, Pierre-Philippe and Laurent Gobillon (2015). The Empirics of Agglomeration Economies. In: Duranton, Gilles, J. Vernon Henderson and William Strange (eds.), Handbook of Regional and Urban Economics 5A, Elsevier, 247–348.

Combes, Pierre-Philippe, Gilles Duranton and Laurent Gobillon (2019). The costs of agglomeration: House and land prices in French cities. Review of Economic Studies 86:4, 1556–1589.

Dauth, Wolfgang and Johann Eppelsheimer (2020). Preparing the Sample of Integrated Labour Market Biographies (SIAB) for Scientific Analysis: A Guide, Journal for Labour Market Research. 54:10.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum (2022). Matching in Cities. Journal of the European Economic Association 20:4, 1478-1521.

Davis, Morris A., Jonas D. M. Fisher and Toni M. Whited (2014). Macroeconomic implications of agglomeration. Econometrica 82:2, 731–764.

De la Roca, Jorge and Diego Puga (2017). Learning by working in big cities. Review of Economic Studies 84:1, 106–142.

Desmet, Klaus and Esteban Rossi-Hansberg (2009). Urban accounting and welfare. American Economic Review 103:6, 2296–2327.

Desmet, Klaus and J. Vernon Henderson (2015). The Geography of Development within Countries, in: Duranton, Gilles, J. Vernon Henderson, and William Strange (eds.), Handbook of Regional and Urban Economics 5B. Elsevier, 1457– 517.

Destatis (2021). Inlandsproduktberechnung - Lange Reihen ab 1970, 18:1.4, Federal Statistical Office of Germany (Statistisches Bundesamt).

Diamond, Rebecca (2016). The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000. American Economic Review 106:3, 479-524.

Diamond, Rebecca and Cecile Gaubert (2022). Spatial Sorting and Inequality. Annual Review of Economics 14, 795-819.

Diamond, Rebecca and Enrico Moretti (2022). Where is Standard of Living the Highest? Local Prices and the Geography of Consumption. NBER Working Paper 16894.

Duranton, Gilles and Diego Puga (2004). Micro-foundations of urban agglomeration economies, Handbook of Regional and Urban Economics 4, J.V. Henderson and J.-F. Thisse (eds.), Amsterdam: Elsevier, 2063–2117.

Duranton, Gilles and Diego Puga (2014). The growth of cities. In: P. Aghion and S. Durlauf (eds.), Handbook of Economic Growth 2. Elsevier, 468–560.

Duranton, Gilles and Diego Puga (2015). Urban Land Use. In: Duranton, Gilles, J. Vernon Henderson, and William Strange (eds.), Handbook of Regional and Urban Economics 5A, Elsevier, 467-560.

Duranton, Gilles and Diego Puga (2019). Urban Growth and Its Aggregate Implications. NBER Working Paper 26591.

Duranton, Gilles and Diego Puga (2020). The Economics of Urban Density. Journal of Economic Perspectives 34:3, 3–26.

Duranton, Gilles and Diego Puga (2022). Urban Growth and Its Aggregate Implications. Revised version of NBER Working Paper 26591.

Duranton, Gilles and Matthew A. Turner (2012). Urban Growth and Transportation. Review of Economic Studies 79:4, 1407-1440.

Dustmann, Christian, Bernd Fitzenberger and Markus Zimmermann (2022). Housing Expenditures and Income Inequality. The Economic Journal 132, 1709-1736.

Eberle, Johanna and Alexandra Schmucker (2019). Creating cross-sectional data and biographical variables with the Sample of Integrated Labour Market Biographies 1975-2014 - programming examples for Stat. Institute for Employment Research, FDZ-Methodenreport 04/2019.

Economist (2015). German property. Braking bad. April 4.

Economist (2017). Poor and sexy. Why is Berlin so dysfunctional? December 2.

Economist (2018). Peak Valley. Why start-ups are going elsewhere. September 1.

Eeckhout, Jan (2004). Gibrat's Law for (All) Cities. American Economic Review 94:5, 1429–1451.

Ethier, Wilfred (1982). National and international returns to scale in the modern theory of international trade. American Economic Review 72, 389–405.

Fajgelbaum, Pablo D., and Cecile Gaubert (2020). Optimal Spatial Policies, Geography and Sorting. Quarterly Journal of Economics 135:2, 959–1036.

Fajgelbaum, Pablo D., Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar (2019). State Taxes and Spatial Misallocation. Review of Economic Studies 86:1, 333–76.

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015). The Next Generation of the Penn World Table. American Economic Review 105:10, 3150-3182.

Fischel, William A. (2001). The Homevoter Hypothesis. Cambridge MA, Harvard University Press.

Follmer, Robert and Dana Gruschwitz (2019). Mobility in Germany – short report. Edition 4.0 of the study by infas, DLR, IVT and infas 360 on behalf of the Federal Ministry of Transport and Digital Infrastructure (BMVI) (FE no. 70.904/15). Bonn, Berlin.

Fujita, Masahisa and Jacques-François Thisse (2013). Economics of Agglomeration: Cities, Industrial Location, and Regional Growth. Cambridge: Cambridge University Press. Second Edition.

Gabaix, Xavier (1999). Zipf's law for cities: An explanation. Quarterly Journal of Economics 114:3, 739-767.

Gabaix, Xavier (2009). Power Laws in Economics and Finance. Annual Review of Economics 1, 255-293.

Gabaix, Xavier and Rustan Ibragimov (2011). Rank-1/2: A simple way to improve OLS estimation of tail exponents. Journal of Business Economics and Statistics 29:1, 24-39.

Gyourko, Joseph, Jonathan S. Hartley and Jacob Krimmel (2021). The local residential land use regulatory environment across US housing markets: Evidence form a new Wharton index. Journal of Urban Economics 124, 103337.

Gyourko, Joseph, Albert Saiz and Anita A. Summers (2008). A new measure of the local regulatory environment for markets: The Wharton Residential Land Use Regulatory Index. Urban Studies 45:3, 693-729.

Gartner, Hermann (2005). The imputation of wages above the contribution limit with the German IAB employment sample. Institute for Employment Research, FDZ Methodenreport 02/2005.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer, Andrei (2013). Human capital and regional development. Quarterly Journal of Economics 128:1, 105-164.

Giesen, Kristian and Jens Suedekum (2011). Zipf's law for cities in the regions and the country. Journal of Economic Geography 11:4, 667–686.

Giesen, Kristian, Arndt Zimmermann and Jens Suedekum (2010). The size distribution across all cities – Double Pareto lognormal strikes. Journal of Urban Economics 68:2, 129-137.

Glaeser, Edward L. (2011). Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier, London: MacMillan.

Glaeser, Edward L. and David Cutler (2021). Survival of the City. Living and Thriving in an Age of Isolation. Penguin Random House.

Glaeser, Edward L. and Joseph Gyourko (2005). Urban decline and durable housing. Journal of Political Economy 113: 2, 345–375.

Grujovice, Anja (2021). Task, Cities and Urban Wage Premia. Mimeo. CEMFI Madrid.

Henderson, J. Vernon (1974). The sizes and types of cities. American Economic Review 64, 640–656.

Henkel, Marcel, Tobias Seidel and Jens Suedekum (2021). Fiscal Transfers in the Spatial Economy. American Economic Journal: Economic Policy 13:4, 433–468

Hsieh, Chang-Tai and Enrico Moretti (2019). Housing constraints and spatial misallocation. American Economic Journal: Macroeconomics 11:2, 1-39.

Jacobs, Jane (1969). The economy of cities. Random House, New York.

Jacobs, Jane (1984). Cities and the wealth of nations. Random House, New York.

Jones, Charles I. (1995). R&D-Based Models of Economic Growth. Journal of Political Economy 103:4, 759–784.

Jones, Charles I. (2002). Sources of U.S. Economic Growth in a World of Ideas. American Economic Review 92:1, 220–239.

Jones, Charles I. (2005). Growth and Ideas. In: Philippe Aghion and Steven A. Durlauf, eds., Handbook of Economic Growth, New York: North Holland, 1063–1111.

Jones, Charles I. (2022). The Past and Future of Economic Growth: A Semi-Endogenous Perspective. Annual Reviews in Economics 14: 125-152.

Klick, Larissa and Sandra Schaffner (2021). Regional Real Estate Price Indices for Germany, 2008–2019: RWI-GEO-REDX. Journal of Economics and Statistics 241:1, 119–129.

Lucas, Robert (1988). On the Mechanics of Economic Development. Journal of Monetary Economics 22 3-42.

Moretti, Enrico (2004a). Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. Journal of Econometrics 121:1, 175-212.

Moretti, Enrico (2004b). Human capital externalities in cities. In: Henderson, J. Vernon and Jacques-Francois Thisse (eds.), Handbook of Regional and Urban Economics 4, Elsevier, 2243-2291.

Moretti, Enrico (2013). Real Wage Inequality. American Economic Journal: Applied Economics 5:1, 65-103.

Nobis, Claudia and Katja Köhler (2018). Mobilität in Deutschland – MiD Nutzerhandbuch, Studie von infas, DLR, IVT und infas 360 im Auftrag des Bundesministers für Verkehr und digitale Infrastruktur.

Pestel-Institut (2023). Bauen und Wohnen in der Krise, Hannover.

Pflüger, Michael (2021). City Size, Pollution and Emission Policies. Journal of Urban Economics 126, 103391.

Postbank (2022). Wohnatlas 2021. Bonn.

PricewaterhouseCoopers (2019). Investing in German Real Estate. Frankfurt and Berlin.

Proost, Stef and Jacques-Francois Thisse (2019). What Can Be Learned from Spatial Economics? Journal of Economic Literature 57:3, 575-643.

Redding, Stephen J. (2020). Trade and Geography. NBER Working Paper 27821.

Redding, Stephen J. and Esteban Rossi-Hansberg (2017). Quantitative Spatial Economics. Annual Review of Economics 9, 21-58.

RWI (2019). RWI-GEO-RED. RWI Real Estate Data (Scientific Use File) - apartments for rent, RWI–Leibniz Institute for Economic Research, Dataset, DOI: 10.7807/immo: red:wm:suf:v1.

RWI and microm (2017a). RWI-GEO-GRID: Socio-Economic Data on grid level (Scientific Use File 5.1) – Purchasing power, RWI–Leibniz Institute for Economic Research, Dataset, DOI: 10.7807/microm:kaufkraft:suf:v5:1.

RWI and microm (2017b). RWI-GEO-GRID: Socio-Economic Data on grid level (Scientific Use File 5.1) – Unemployment rate, RWI–Leibniz Institute for Economic Research, Dataset, DOI: 10.7807/microm:alq:suf:v5:1.

Schaffner, Sandra (2020). FDZ Data description: Real-Estate Data for Germany (RWI-GEORED v3) - Advertisements on the Internet Platform ImmobilienScout24 2007-06/2020, RWI Projektberichte.

Turner, Matthew, Andrew Haughwout, and Wilbert van der Klaauw (2014). Land Use Regulation and Welfare. Econometrica 82:4, 1341-1403.

Valentinyi, Ákos and Berthold Herrendorf (2008). Measuring factor income shares at the sectoral level. Review of Economic Dynamics 11:4, 820-835.

Weinand, Sebastian and Ludwig von Auer (2020), Anatomy of Regional Price Differentials: Evidence From Micro Price Data. Spatial Economic Analysis 15:4, 413-440.

Appendix: The Worth of Cities in Germany

by Katja Gehr and Michael Pflüger, May 9, 2023

A Data Appendix

A.1 BBSR-urban district regions

Definition of district regions. In order to keep consistency with our labor market data taken from the SIAB-R dataset (Sample of Integrated Labour Market Biographies, Regional File), our definition of cities draws on the classification of district regions by the Federal Institute for Research on Building, Urban Affairs, and Spatial Development (BBSR) at the territorial allocation as on December 31, 2017. The district regions are constructed from the autonomous municipal authorities (*kreisfreie Städte*) and administrative districts (*Kreise*) of Germany such that each district region has a population above 100,000 inhabitants in the reference year 2017. This yields in total 328 district regions (Antoni et al., 2019: 24). We further aggregate the district regions whenever a city is completely enclosed by its county and, therefore, the city constitutes the center of the county.[38] This results in 315 aggregated district regions that are displayed in panel A of fig. A1.

Panel A: Territorial boundaries   Panel B: Urban and rural areas

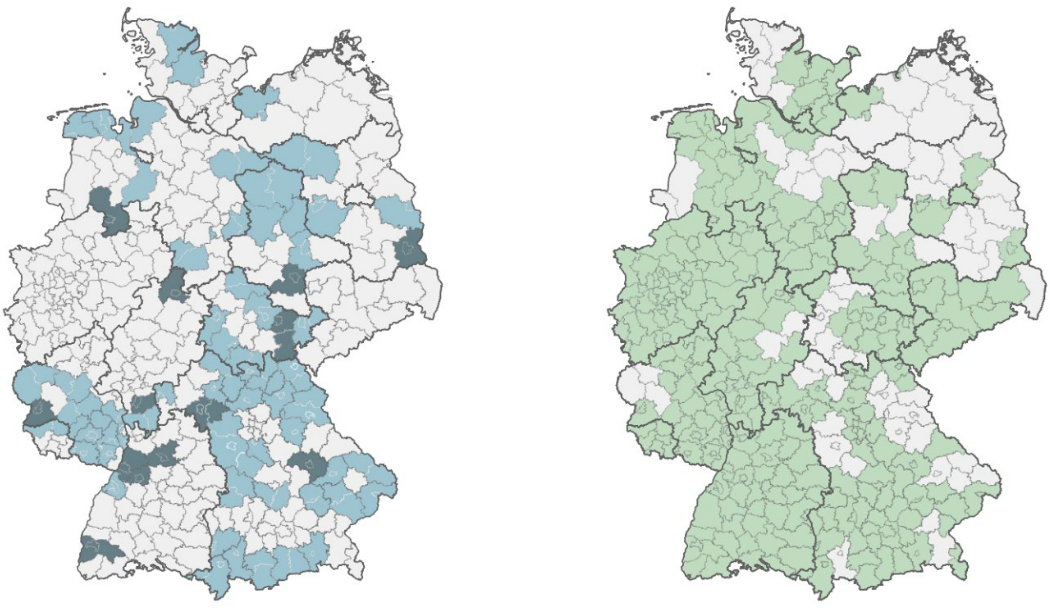

Figure A1: Aggregated district regions

Notes: Panel A visualizes the aggregation of German autonomous municipal authorities and administrative districts to district regions. Light blue areas indicate districts that are summarized to district regions such that each region has a population above 100,000 inhabitants. Dark blue areas indicate the further aggregation of district regions whenever a city is surrounded by its county. Panel B shows the spatial distribution of urban and rural areas. Green/gray areas indicate district regions classified as urban/rural.

[38] Precisely, this applies to the following district regions: Darmstadt and Darmstadt-Dieburg, Spree-Neiße and Cottbus, Saalekreis and Halle (Saale), Freiburg im Breisgau and Breisgau-Hochschwarzwald, Heilbronn and the city of Heilbronn, Trier-Saarburg and Trier, Regensburg and the city of Regensburg, the city of Würzburg and Würzburg/Kitzingen, Enzkreis and Pforzheim, Karlsruhe and the city of Karlsruhe, Osnabrück and the city of Osnabrück, Kassel and the city of Kassel, Jena and Saale-Holzland-Kreis/Saale-Orla-Kreis.

Definition of cities. The aggregated district regions are assigned to urban and rural areas by the classification by the Federal Institute for Research on Building, Urban Affairs, and Spatial Development (BBSR), see BBSR 2012. This classification defines four district types and three region types by the degree of urbanization. The district types are constructed on the following criteria: share of population in large and medium-sized cities, population density, and population density without taking large and medium-sized cities into account. The region types are based on the criteria: share of population in large and medium-sized cities, presence and size of a large city, population density, and population density without taking large cities into account. The precise district and region types with the corresponding differentiation criteria can be found in tables A1 and A2.

Table A1: Differentiation criteria for district types

District-free cities	- district-free cities with at least 100,00 inhabitants
Urban districts	- districts with a population share in large and medium-sized cities of at least 50% and a population density of at least 150 inhabitants per km^2 - districts with a population density outside large and medium-sized cities of at least 150 inhabitants per km^2
Rural districts with beginning agglomeration	- districts with a population share in large and medium-sized cities of at least 50%, but with a population density of less than 150 inhabitants per km^2 - districts with a population share in large and medium-sized cities of less than 50% and a population density outside large and medium-sized cities of at least 100 inhabitants per km^2
Sparsely populated rural districts	- districts with a population share in large and medium-sized cities below 50% and a population density outside large and medium-sized cities of less than 100 inhabitants per km^2

We classify the aggregated district regions, that are assigned to sparsely populated rural districts in rural regions by the BBSR, as rural areas. The remaining aggregated district regions are classified as urban areas and we take them as corresponding to the cities in the model. Therefore, we classify 264 of the 315 aggregated district regions as urban. This yields a share of population living in urban areas in 2017 of 89.17% and a share of territorial area declared urban of 69.79%. The assignment of the aggregated district regions to urban and rural areas is illustrated in panel B of fig. A1.

Table A2: Differentiation criteria for region types

Urban regions	- regions with at least 50% of the population living in large and medium-sized cities and with a large city of at least 500,000 inhabitants - regions with a population density outside large cities of at least 300 inhabitants per km^2
Regions with inter-mediate urbanization	- regions with at least 33% of the population living in large and medium-sized cities and with a population density between 150 and 300 inhabitants per km^2 - regions with at least one large city and a population density outside large cities of at least 100 inhabitants per km^2
Rural regions	- regions with less than 33% of the population living in large and medium-sized cities and with a population density of less than 150 inhabitants per km^2 - regions with a large city, but a population density outside large cities of less than 100 inhabitants per km^2

Definition of city centers. Following the approach of Duranton and Puga (2019), we define the city center of a district region as the location indicated by Google Maps for the core of the main city of the district region. We then calculate the distance to the city center as the haversine distance, which takes the curvature of the earth's surface into account, between the midpoint of each grid cell and the center of the district region.

Population data for 1995 and 2017. Our population data for the administrative districts in Germany for the year 2017 are taken from 'Regionaldatenbank Deutschland' provided by the Federal Statistical Office of Germany. The population levels are in consistency with the territorial boundaries of the administrative districts on 31 December 2017. The population data is then aggregated to district regions.

Due to several district reforms in Germany between 1995 and 2017, the territorial boundaries of district regions in 1995 were substantially different from the boundaries in 2017. Hence, data for the population of the district regions in 1995 with the territorial boundaries as on 31 December 2017 is required for comparability between the years 1995 and 2017. As the population dataset from 'Regionaldatenbank Deutschland' for the year 1995 corresponds to the territorial boundaries of the administrative districts on 31 December 1995, this data source is not suitable for our analysis. Instead, population data from 'Laufende Raumbeobachtung des BBSR' provided by the BBSR is used. As described in BBSR (2010), the territorial boundaries of the administrative districts are assigned to the status as on 31 December 2017. Moreover, the dataset is adjusted for the census of 2011 (BBSR, 2016). The resulting population dataset provides a time consistent definition for the territorial boundaries of the German districts such that the population of district regions in 1995 and 2017 can directly be compared.

A.2 Sample of Integrated Labour Market Biographies (SIAB-R)

Our estimation of the benefits of urban agglomeration is based on the factually anonymous Sample of Integrated Labour Market Biographies (SIAB) of the Institute for Employment Research (IAB) (version 1975 – 2017). The regional file of the SIAB is provided by the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the IAB. The dataset contains a 2% random sample of the administrative social security records from 1975 to 2017 including employees subject to social security contributions and marginal part-time employment. Hereby, the information on employment histories comes from the Employee History (Beschäftigtenhistorik - BeH). The SIAB dataset comprises employment biographies for 1,827,903 individuals with a total of 62,340,521 observations (Antoni et al., 2019: 6). Because of anonymization, the regional file of the SIAB dataset includes 47,536 individuals less than the original SIAB dataset such that 2.5% of the individuals are not contained in the regional file (Antoni et al., 2019: 25). The dataset is in spell format where the unit of observation is any change in the employment status of an individual. The workplace of an individual is observable at the level of district regions. District regions are constructed from the administrative districts of Germany by aggregation to at least 100,000 inhabitants per district region. This gives in total 328 district regions (Antoni et al., 2019: 24). Further details on the district regions can be found in appendix A.1.

Following Dauth and Eppelsheimer (2020), we first convert the dataset into a yearly panel with June 30 as the cut-off date. This date is chosen as the variables on establishments from the Establishment-History-Panel (Betriebs-Historik-Panel - BHP) are not spell data but are only exact on 30 June each year. Furthermore, we restrict the dataset to the time period from 1993 to 2017 since the social security notifications in East Germany are only complete from 1993 on because of the German reunification in 1990 (Antoni et al., 2019: 21). Fig. A2 plots the distribution of wages for the year 2017 in SIAB-R dataset. As can be seen in the figure, the bar furthest to the right is exceptionally high. This corresponds to the top-censoring of wages above the upper earnings limit for statutory pension insurance in the dataset. The wages that exceed the upper earnings limit are not excluded from the dataset but are set equal to the upper earnings limit (Antoni et al, 2019: 42). This yields the high bar at the furthest right of fig. A2. Since the procedure of top censoring of wages would bias our estimation, wages corresponding to the upper earnings limit for statutory pension insurance are dropped from the sample.

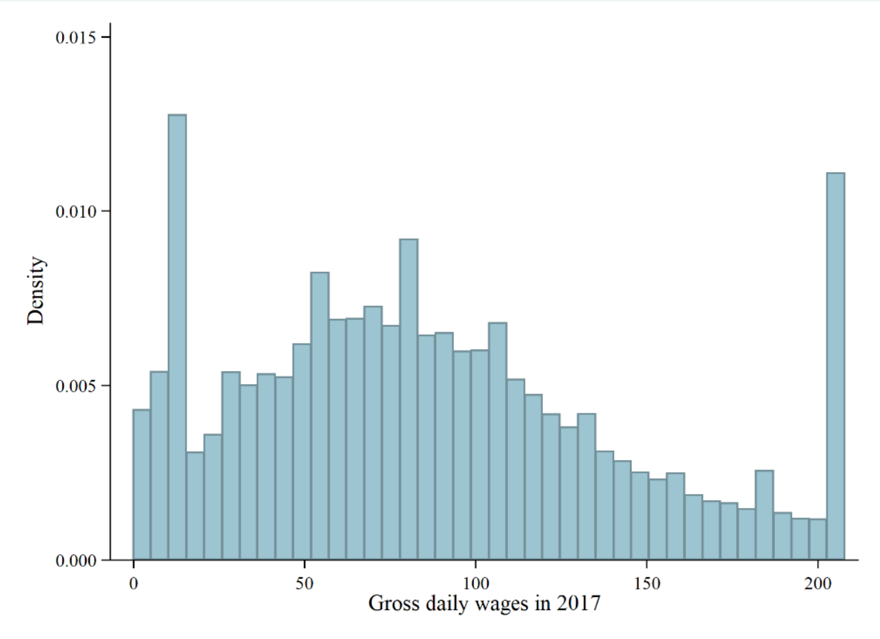

Figure A2: Distribution of wages in 2017

Notes: The figure reports the distribution of the censored daily wage for the year 2017 using the SIAB-R data.

Moreover, we constrain the sample to full-time employees liable to social security because only the daily wage but not the underlying hours worked are observed for part-time workers. Hence, it is not feasible to control for changes in part-time workers' wages due to changes in the underlying hours worked. As marginal employments are only included since 2011 in the dataset, we also exclude these observations from the sample. Further, we only include workers of German nationality since the full labor market histories cannot be observed for foreign born workers. To constrain the sample to workers at least 18 years old for whom the full labor market histories are observable, we restrict the year of birth to years between 1975 and 1999. Finally, we drop all observations for which no workplace location is observed from the sample and restrict the dataset to urban areas as defined in appendix A.1. This yields a final sample of 1,529,134 observations.

The variables work experience and firm tenure, included as controls in our estimation of the urban benefits, are constructed following Eberle and Schmucker (2019). The variable firm tenure measures the number of days an individual was employed in an establishment expressed in years. Training periods in the establishment are included in our measure of firm tenure. The variable work experience counts the number of days an individual has been employed up to the current point in time and is measured in years. Training periods are excluded from our measure of work experience. Descriptive statistics on the variables used in the estimation of the urban benefits can be found in table A3.

Table A3: Descriptive statistics on the SIAB-R data

Variable	Mean	St. Error	1st decile	Median	9th decile
Gross daily wage (in €)	89.35	34.84	48.76	84.82	138.31
Work experience	6.24	4.42	1.13	5.39	12.68
Firm tenure	3.90	3.82	0.33	2.82	9.24
Experience in five biggest cities	0.70	2.19	0.00	0.00	2.28
Experience in cities >500,000 (without five biggest)	0.72	2.25	0.00	0.00	2.37
Very-high-skilled occupations	0.11	0.31	0.00	0.00	1.00
High-skilled-occupations	0.10	0.30	0.00	0.00	1.00
Medium-skilled occupations	0.73	0.45	0.00	1.00	1.00
Low-skilled occupations	0.06	0.25	0.00	0.00	1.00

A.3 Real-estate data (RWI-GEO-RED)

Our estimation of the urban cost parameter, γ, uses rent prices for Germany coming from the real-estate RWI-GEO-RED data provided by the Research Data Centre Ruhr at RWI (RWI, 2019). The dataset is based on real estate advertisements from the largest German listing platform ImmobilienScout24. It includes information on prices and various further characteristics of apartments as well as houses for sale and for rent. These further characteristics cover size, facilities and equipment, additional costs, and power consumption of a real estate. As the advertised price of an object is not binding, the prices in the dataset correspond to offering prices at which an owner is willing to sell or rent an object. The dataset is provided on a monthly basis for the years 2007 to 2020. The location of a real estate is provided on a grid of 1-square-kilometre raster cells covering whole Germany. The projection on the 1-square-kilometre grid raster follows the European standard ETRS89-LAEA according to the INSPIRE guidelines. Each grid cell is assigned to an administrative district in the dataset (Schaffner, 2020).

We prepare the dataset as follows for our analysis. We restrict the dataset to apartments for rent advertised in the year 2017. The year 2017 is chosen to keep consistency with the datasets on commuting behavior (MiD 2017) and on labor market histories (SIAB-R). Since the territorial status assigned to the administrative districts is as on 31 December 2015, we reassign grid cells to administrative districts based on territorial definitions on 31 December 2017. Then, we aggregate the administrative districts to district regions as described in appendix A.1 and restrict the dataset to apartments located in district regions classified as urban. In the dataset, the problem arises that some advertisement identifiers are not unique as described in Schaffner (2020). This can occur for the following main two reasons. First, an advertisement was not concluded at the time of data delivery and, therefore, is also included in the next data delivery. The advertisement is then included twice in the dataset. Second, an old advertisement is used as a template for a new advertisement. Then, the same advertisement identifier corresponds to different objects. Following the guideline of Schaffner (2020), we drop apartments from the data if objects with the same identifier are classified as similar based on their observed characteristics and if in addition, the gap between the advertisements is not larger than six months. Following Klick and Schaffner (2019), we drop apartments with a rent above 5,000€

Table A4: Descriptive statistics on the RIW-GEO-RED data

Variable	Mean	St. Error	1st decile	Median	9th decile
Rent (€ per m^2)	8.67	4.06	5.01	7.67	13.33
Dwelling area (€ per m^2)	71.47	29.31	40.00	66.00	107.13
Number of rooms	2.51	0.96	1.00	2.00	4.00
Construction decade	1960	40.72	1900	1970	2010
Object categories					
top floor apartment	0.17	0.38	0.00	0.00	1.00
apartment	0.66	0.47	0.00	1.00	0.00
mezzanine	0.03	0.17	0.00	0.00	1.00
maisonette	0.01	0.07	0.00	0.00	0.00
penthouse	0.05	0.22	0.00	0.00	0.00
basement	0.01	0.12	0.00	0.00	0.00
terrace apartment	0.05	0.22	0.00	0.00	0.00
other	0.01	0.11	0.00	0.00	0.00
Distance from grid cell to city center (in km)	7.36	7.99	1.01	4.40	17.92
Number of grid cells per urban area	166.13	90.80	68	153	273
Number of observations per urban area	3531	6278	670	1862	6783

or a living area exceeding 400 square meters from the sample as these apartments are very luxurious and not representative for the German real estate market. Furthermore, we exclude apartments with a rent below 50€ or a living area smaller than 50 square meters from the sample. These advertisements mainly correspond to parking spaces, cellar compartments, and workrooms and, hence, do not belong to rental apartments in a narrow sense. We construct the variable rent per square meter used in our estimation by dividing the exclusive rent of an apartment by its living area.

After our outlined data preparation steps, a final sample with 906,315 observations is obtained. The sample covers all 264 urban district regions with an average of 166 grid cells per district region. This yields, on average, 3531 observations per urban district region. The mean rent in the sample is 8.67€ with a mean dwelling area of 71.47 sqm. Further descriptive statistics on the main variables in the dataset can be found in table A4.

A.4 Mobility in Germany (MiD)

Our estimation of the elasticity of congestions costs draws on the dataset 'Mobilität in Deutschland (MiD)' which is a nationwide survey instructed by the Federal Ministry of Transport and Digital Infrastructure in Germany (BMVI) provided by the infas Institute for Applied Social Sciences (BMVI, 2017). The survey is conducted in the year 2017. Households are selected randomly and questioned on their socio demographic background and their travel behavior on a given reference date. On the household level, the dataset contains information on the household size, age structure, economic status, place of residence, and endowment with means of transport. Age, sex, common use of transport, and the possession of a driver's license are included on the individual level in the dataset. For the travel behavior on the survey day, departure and arrival time, commuting distance, trip purpose, use of transport, and the travel speed are most importantly reported for our analysis. The place of residence of an individual is observed on a grid of 1-square-kilometre raster cells covering all parts of Germany. The projection on the 1-square-kilometre raster grid follows the European standard ETRS89-LAEA according to INSPIRE. As grid cells are not assigned to German administrative districts in the dataset, this is done as part of the data preparation. We allocate each grid cell to a district region according to the largest share of surface area of a district region that is included in a grid cell. (Follmer and Gruschwitz, 2019)

The dataset contains a total sample of 316,361 individuals from 156,420 households with 960,619 reported commuting trips. We restrict the sample to commuting trips by private car in order to estimate the urban cost parameter, θ, only through congestion that arises from traffic jams. Further, an interviewed person is only included if the person drove herself. Otherwise, the possibility arises that an individual is passenger in another trip reported, for example, by a household member of the interviewed person. Then the trip would be included twice in the sample: once for the driver and once for the passenger of the same trip. Moreover, a trip is only included in the sample if the interviewed person is in her usual environment on the reference date such that the commuting tip is assigned to the correct district region of the driver's place of residence. Finally, the sample is restricted to drivers whose place of residence is located in an urban area to keep consistency with the theoretical model. This yields a sample of 57,034 individuals from 48,161 households with a total of 187,435 commuting trips. Descriptive statistics on the main variables in the dataset are contained in table A5.

Table A5: Descriptive statistics on the MiD 2017 data

Variable	Mean	St. Error	1st decile	Median	9th decile
Travel speed (in km/h)	29.41	19.67	9.52	24.43	57.00
Travel distance (in km)	15.79	40.97	1.43	5.70	31.35
Retire	0.27	0.44	0.00	0.00	1.00
Age groups					
18 – 29 years	0.07	0.27	0.00	0.00	0.00
30 – 39 years	0.10	0.30	0.00	0.00	1.00
40 – 49 years	0.18	0.38	0.00	0.00	1.00
50 – 59 years	0.27	0.44	0.00	0.00	1.00
60 – 69 years	0.20	0.40	0.00	0.00	1.00
70 – 79 years	0.14	0.34	0.00	0.00	1.00
80 years and older	0.03	0.18	0.00	0.00	0.00
Trip purpose					
Commute to workplace	0.18	0.38	0.00	0.00	1.00
On business	0.13	0.34	0.00	0.00	1.00
Commute to education place	0.01	0.08	0.00	0.00	1.00
Leisure	0.22	0.41	0.00	0.00	1.00
Grocery shopping	0.21	0.40	0.00	0.00	1.00
Errands	0.16	0.37	0.00	0.00	1.00
Distance from grid cell to city center (in km)	9.82	8.50	1.66	7.37	21.02
Number of observations per urban area	926	1299	127	573	1883

A.5 Geographical constraints

To quantify the share of developable land Λ_i, we follow Saiz (2010) and calculate for each city the share of geographically unconstrained land within the city's 30-km radius of the city center. Hereby, an area is classified as geographically constrained if it is covered by water and wetlands, slopes steeper than 15%, nature reserves or if it belongs to a foreign country.

The slope is calculated using the EU Digital Elevation Model at its 25-meter resolution from the Copernicus Land Monitoring Service and all areas with slopes steeper than 15% are classified as geographically constrained. Data on water and wetlands including seas, lakes, rivers, and other internal water bodies comes from the CORINE Land Cover 5 ha CLC5 for the year 2018. Nature reserves are identified with OpenStreetMapData and foreign land is identified from the official boundary files of the Federal Agency for Cartography and Geodesy.

B Robustness check – Population density

For our estimations of the urban benefits and costs, it must be decided whether a city's population or density should be used in the regressions to measure the size of a city. On the one hand side, cities with a high density but small population size are unlikely to exhibit strong agglomeration economies. On the other hand side, workers at the edge of a large district region where density is low are unlikely to fully benefit from agglomeration economies through human capital spillovers. Hence, neither a city's density nor population size alone can fully account for the agglomeration effects in these cases. Empirically, a positive correlation between city density and city population can be found such that both measurements used in a regression should result up to a certain point in similar results (Duranton and Puga, 2020: 7).

Table B1: 20 largest cities by population and density

rank	city	population	rank	city	density
1.	Berlin	3,613,495	1.	Munich	4686
2.	Hamburg	1,830,584	2.	Berlin	4055
3.	Munich	1,456,039	3.	Stuttgart	3052
4.	Region of Hannover	1,152,675	4.	Herne	3043
5.	Cologne	1,080,394	5.	Frankfurt am Main	3008
6.	Karlsruhe	754,592	6.	Düsseldorf	2839
7.	Frankfurt am Main	746,878	7.	Offenbach am Main	2822
8.	Stuttgart	632,743	8.	Essen	2774
9.	Düsseldorf	617,280	9.	Nuremberg	2763
10.	Recklinghausen	616,824	10.	Oberhausen	2743
11.	Rhein-Sieg-Kreis	599,056	11.	Cologne	2668
12.	Dortmund	586,600	12.	Bochum	2509
13.	Essen	583,393	13.	Gelsenkirchen	2481
14.	Leipzig	581,980	14.	Hamburg	2424
15.	Bremen	568,006	15.	Bonn	2307
16.	Region of Aachen	554,068	16.	Mainz	2201
17.	Dresden	551,072	17.	Ludwigshafen am Rhein	2176
18.	Rhein-Neckar-Kreis	546,745	18.	Duisburg	2140
19.	Ludwigsburg	546,745	19.	Mannheim	2125
20.	Esslingen	532,447	20.	Wuppertal	2100

However, not only theoretical aspects should be taken into account but also practical considerations. Table B1 displays the twenty largest district regions in Germany respectively ranked by their population size and density. Hereby, we calculate the density of a district region by dividing the population of a district region by its area. As the regional unit underlying the analysis is the district region, distortions arise in the ranking of the size of the district regions. From table B1 can be taken that some district regions with relatively low expected agglomeration economies are assigned with a relatively large city size. For example, the district regions Recklinghausen and Rhein-Sieg-Kreis have in aggregate a very large population size. However, these district regions are composed of many small cities and lack the presence of a bigger city with strong agglomerations. In the case of density, this argumentation for example applies to the district regions Herne and Oberhausen. In these cases, the borders of the district regions are placed relatively close to the city limits, whereas in other district regions the surrounding countryside of a city is also included in the boundaries of the district region. As a result, the density is overestimated in some cities. As the described distortions seem to be stronger under the consideration of the density of a district region, we decide for using the district region's population as the measurement of its size in our baseline specification. In addition, however, the alternative regression results based on the city's density are shown in tables B2 and B3.

Table B2: Estimation of elasticity of congestion costs/travel speed based on density

Dependent variable:	Ln estimated city travel speed
Ln density	-0.0394^{***}
	(0.0038)
Observations	264
R^2	0.3343

Notes: The regression includes a constant term. City travel speed is estimated in a previous step by regressing travel speed for individual trips by private car on city indicators, including grid cell, driver and trip controls. We use this to predict for each city the speed of a 15km commuting trip on a Tuesday at 8AM by a driver with average characteristics. Coefficients are reported with robust standard errors in parenthesis. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Table B3: Estimation of population elasticities of urban benefits based on density

	(1)	(2)	(3)	(4)
Dependent variable:	Ln earnings		Static premium (city indicators column (2))	Medium-term premium (static + 6.3 years local experience)
Ln city density	0.0206***		0.0290***	0.0357***
	(0.0007)		(0.0044)	(0.0046)
City indicators		Yes		
Worker fixed effects	Yes	Yes		
Experience in five biggest cities	0.0135***	0.0140***		
	(0.0008)	(0.0008)		
Experience in five biggest cities x exp.	-0.0006***	-0.0006***		
	(0.0001)	(0.0001)		
Experience in cities > 500,000 (without five biggest)	0.0100***	0.0100***		
	(0.0007)	(0.0010)		
Experience in cities > 500,000 (without five biggest) x exp.	-0.0004***	-0.0004***		
	(0.00001)	(0.0001)		
Experience	0.0468***	0.0462***		
	(0.0007)	(0.0007)		
Experience^2	-0.0016***	-0.0016***		
	(0.0001)	(0.0001)		
Observations	1,529,134	1,529,134	264	264
R^2	0.3294	0.3382	0.1430	0.1959

Notes: All regressions include a constant term. Columns (1) and (2) include firm tenure and its square, year indicators, 4 occupational skill indicators, 14 sector indicators, and 120 occupation indicators. Column (2) in addition includes 264 city indicators. Worker values of experience and tenure are calculated on the basis of actual days worked and expressed in years. Coefficients are reported with robust standard errors in parenthesis, which are clustered by worker in columns (1) and (2). ***, **, and * indicate significance at the 1, 5, and 10 percent levels. The R^2 reported in columns (1) and (2) is within workers.

C Quantification and algorithms for the counterfactuals

In this part of the appendix, we describe the quantification and algorithm that is used to calculate the counterfactual scenarios.

C.1 Quantification

The empirical city sizes in 2017 are taken as the equilibrium with local governments in the model, i.e., factual city sizes in 2017 correspond to N_{it}^*. We follow Duranton and Puga (2022) in assuming that the city populations in the first year from which our analysis starts, i.e., the year 1995, are the cities' incumbents and we take the additional city population that enters cities from 1995 to 2017 to be the newcomers (if positive). We employ the following approximations for the estimated parameters: $\sigma = 0.017$, $\eta = 0.020$, $\gamma = 0.071$ and $\theta = 0.068$, and we use $\lambda = 0.18$.

Cities. The productive amenities are filtered from the wage equation, where w_{it} is approximated with the average annual wage in the city for the year 2017

$$A_{it}^p B(\psi)^{1+\sigma} = w_{it} / {N_{it}^*}^{\sigma+\eta}. \tag{C1}$$

The fiscal transfer rate is calculated from the optimality condition for the urban costs by normalizing $\tau_t \equiv 10{,}000$ as

$$\varphi_i = \frac{net{-}transfers_{it-HSS}}{nuc_{it} \cdot N_{it}} = \frac{(1+\gamma){\Lambda_i}^{\gamma} \cdot net{-}transfers_{it-HSS}}{\tau_t \, {N_{it}^*}^{\gamma+\theta+1}}. \tag{C2}$$

This calculation uses the net-transfers across German cities, $net - transfers_{it-HSS}$, which were established by Henkel et al. (2019) for their own project and which they generously provided to us (see footnote 22 in the main text).

The consumptive amenities can be filtered from the equation for the optimal city size as follows,

$$A_i^c = \frac{\gamma+\theta}{(\sigma+\eta)(1+\gamma)} \frac{(1-\varphi_i)\,\tau_t}{A_{it}^p B(\psi)^{1+\sigma} \Lambda_i^{\gamma}} {N_{it}^*}^{(\gamma+\theta)-(\sigma+\eta)}. \tag{C3}$$

Rural area. To calculate the fiscal transfer per capita in the rural area κ_r, we sum up the transfers across all city locations. The (negative) balance divided by the rural population then captures the transfers per person from cities to the rural area.

The productivity in the rural area is calculated from the condition that the real income in the smallest city has to be equal to the real income in the rural area, $\underline{v}_{it}^* = v_{rt}^*$ as

$$A_{rt} = \left(\underline{v}_{it}^* - \kappa_r\right) {N_{rt}^*}^{\lambda}. \tag{C4}$$

The spatial distribution of productive amenities, consumptive amenities, and fiscal transfer rates is illustrated by figure C1.

Panel A: Productive amenities   Panel B: Consumptive amenities   Panel C: Fiscal transfer rates

Figure C1: Productive amenities, consumptive amenities, and fiscal transfer rates

Notes: The maps visualize the spatial distribution of productive amenities, consumptive amenities, and fiscal transfer rates. In panels A and B, darker shading indicates higher values. In panel C, darker areas indicate recipient regions and bright areas donor regions.

C.2 Counterfactual: Expansion of the Top Seven

The city population is increased by 10% in the following seven cities: Berlin, Hamburg, Munich, Cologne, Frankfurt, Stuttgart, and Düsseldorf ($\widehat{N}_{it} = 1.1N_{it}$). The scenario is implemented by a reduction in land-use regulations in these cities.

Algorithm. Incumbents in other cities than the top seven have no incentive for increasing their cities' housing supply and their city sizes remain unchanged by adjusting their land-use regulations. Hence, the residents migrating to the top seven cities first come from the rural area. Due to the fall in rural population, income in the rural area increases. When the income in the rural area rises, it exceeds the real income of incumbents in the cities with the lowest real incomes. Consequently, the cities with the lowest real incomes are vacated, and the former residents of these cities also migrate to the top seven cities.

Residents from the rural area and the least productive cities migrate to the top seven cities until the size of these cities increases by 10 percent, respectively, and the spatial equilibrium condition is restored. This means that the real income of residents in the new least developed city is equal to the income in the rural area.

New distribution of population. First, the number of residents that move to the top seven cities are calculated such that the city sizes of these seven cities increase by 10 percent, respectively. Then, the following algorithm is executed iteratively.

1. The least developed city is vacated, and the residents move to the top seven cities. The remaining residents, that are still needed to let the city sizes of these seven cities increase by 10 percent, are assumed to migrate from the rural area to those cities.
2. In response to the fall in rural population, the income in rural areas increases. The new income in the rural area is calculated according to

$$\hat{v}_{rt} = A_{rt}\hat{N}_{rt}^{\ -\lambda} + \kappa_r \tag{C5}$$

   and compared to the real income of the new least developed city without land-use regulations.
3. If the real income in the new least developed city is equal to the income in the rural area, the algorithm stops as the spatial equilibrium condition is retained again. If the income in the rural area exceeds the real income in the new marginal city, the algorithm proceeds by in addition vacating the second least productive city.

Real income changes. The real income change for the rural population and newcomers (independently of whether they remain at their location or move to a new city) is given by

$$v_{rt} = A_{rt}N_{rt}^{\ -\lambda} + \kappa_r \qquad \Rightarrow \qquad \frac{\hat{v}_{it}^{\ newc}}{v_{it}^{*\ newc}} = \frac{\hat{v}_{rt}}{v_{rt}^{*}} = \frac{A_{rt}\hat{N}_{rt}^{\ -\lambda}+\kappa_r}{A_{rt}N_{rt}^{*\ -\lambda}+\kappa_r} - 1, \tag{C6}$$

where $\hat{N}_{rt}$ is the counterfactual population in the rural area. The real income loss for incumbents in the top seven cities is given by

$$v_{it} = A_{it}^{p}B(\psi_t)^{1+\sigma}N_{it}^{\ \sigma+\eta} - \frac{1-\varphi_i}{A_i^c\Lambda_i^{\ \gamma}}\frac{\tau_t}{1+\gamma}N_{it}^{\ \gamma+\theta} \qquad \Rightarrow \qquad \frac{\hat{v}_{it}}{v_{it}^{*}} - 1, \tag{C7}$$

where $N_{it} = \hat{N}_{it}$ is used for the calculation of $\hat{v}_{it}$ and $N_{it} = N_{it}^{*}$ for the calculation of v_{it}^{*}. It is assumed that all residents that migrate to the top seven cities become newcomers in these seven cities, independently of whether they have been incumbents or newcomers in their city of origin. Further, the counterfactual real income for newcomers in cities is equal to the counterfactual income in the rural area. Hence, the real income gain for incumbents displaced from city j to one of the top seven cities i is given by

$$\frac{\hat{v}_{rt}}{v_{jt}^{*}} - 1 = \frac{A_{rt}\hat{N}_{rt}^{\ -\lambda}+\kappa_r}{v_{jt}^{*}} - 1, \qquad \text{with } v_{jt}^{*} = \frac{(\gamma+\theta)-(\sigma+\eta)}{(\sigma+\eta)(1+\gamma)}\frac{\tau_t(1+\varphi_i)}{A_i^c\Lambda_i^{\ \gamma}}N_{it}^{*\ \gamma+\theta}. \tag{C8}$$

C.3 Counterfactual: Social optimum

A counterfactual shift from the general equilibrium under the local government allocation to the social optimum city size distribution characterized in section 2.3 is considered.

Algorithm. The optimal city size distribution is determined by the following set of equations:

$$\mu = \underline{v}_{it}^{opt} = \underline{A}_{it}^{p} B(\psi)^{1+\sigma} \underline{N}_{it}^{opt^{\sigma+\eta}} - \frac{1-\varphi_i}{(1+\gamma_i) A_i^c \Lambda_i^{\gamma_i}} \underline{N}_{it}^{opt^{\gamma+\theta}} \tag{C9}$$

$$\mu = (1+\sigma+\eta) B(\psi)^{1+\sigma} A_{it}^{p} N_{it}^{opt^{\sigma+\eta}} - \frac{(1+\gamma+\theta)(1-\varphi_i)}{(1+\gamma) A_i^c \Lambda_i^{\gamma}} N_{it}^{opt^{\gamma+\theta}} \tag{C10}$$

$$N_{rt}^{opt} = \left(\frac{A_{rt}}{\mu-\kappa_r}\right)^{1/\lambda} \tag{C11}$$

and the population constraint $N_{rt}^{opt} + \int_{\underline{A}_{it}^{p}}^{\infty} N_{it}\left(A_{it}^{p}\right) dG\left(A_{it}^{p}\right) = N_t$. As the city with the lowest real income that is developed last in the social optimum is at its locally efficient scale, it holds that $\underline{N}_{it}^{opt} = N_{it}^{*}$.

The optimal city size distribution is calculated iteratively. First, it is assumed that only the two cities with the highest real income are developed in the social optimum. The Lagrange parameter is then pinned down by the real income of the city with the lowest real income (here: the second city) that is developed at its locally efficient size. Given a value for μ, the remaining optimal city sizes (here: the first city) are calculated numerically from the condition:

$$\mu = (1+\sigma+\eta) B(\psi)^{1+\sigma} A_{it}^{p} N_{it}^{opt^{\sigma+\eta}} - \frac{(1+\gamma+\theta)(1-\varphi_i)}{(1+\gamma) A_i^c \Lambda_i^{\gamma}} N_{it}^{opt^{\gamma+\theta}}. \tag{C12}$$

The population in the rural area is calculated for a given value of μ by

$$N_{rt}^{opt} = \left(\frac{A_{rt}}{\mu-\kappa_r}\right)^{1/\lambda}. \tag{C13}$$

Last, it is checked whether the distribution of the population among cities and the rural area add up to the total population in the economy (Germany). If this is the case, the optimal city size distribution is found. If not, the city with the next highest real income is developed.

Real income changes. For the change in real income, it is assumed that all residents that migrate to a new city become newcomers in these cities in the counterfactual scenario. The real income change for the rural population and newcomers (independently of whether they remain at their location or move to a new city) is given by $v_{rt}^{opt}/v_{rt}^{*} - 1 = \frac{A_{rt} N_{rt}^{opt^{-\lambda}} + \kappa_r}{A_{rt} N_{rt}^{*^{-\lambda}} + \kappa_r} - 1$. The real income change for remaining incumbents in cities is given by $v_{it}^{opt}/v_{it}^{*} - 1$. The real income change for incumbents displaced from city j is given by $v_{rt}^{opt}/v_{jt}^{*} - 1$.

C.4 Counterfactual: Fiscal transfers

Under local governments, all fiscal transfers are abolished, i.e., φ_i is set equal to zero for all cities and κ_r is set to zero for the rural area. Then, local governments again install their optimal city sizes by levying land-use regulations on newcomers, but now without the fiscal transfer scheme.

Algorithm. First, the locally optimal city sizes and corresponding real incomes without fiscal transfers under local governments are calculated according to

$$\widehat{N}_{it} = \left(\frac{(\sigma+\eta)(1+\gamma)}{\gamma+\theta} \frac{A_{it}^{p} B(\psi_t)^{1+\sigma} A_i^c {\Lambda_i}^{\gamma}}{\tau_t}\right)^{\frac{1}{(\gamma+\theta)-(\sigma+\eta)}} \text{ and } \hat{v}_{it} = \frac{(\gamma+\theta)-(\sigma+\eta)}{(\sigma+\eta)(1+\gamma)} \frac{\tau_t}{A_i^c {\Lambda_i}^{\gamma}} {\widehat{N}_{it}}^{\gamma+\theta}. \tag{C14}$$

Local governments install the new locally optimal city sizes by adjusting the land-use regulations in place. Cities receiving fiscal transfers become smaller in size when transfers are abolished, and the population of donor cities increases such that residents move from recipient cities to donor cities (the movers are assigned to a city of origin by the initial population shares of the potential cities of origin). As the overall population living in cities is smaller without fiscal transfers, part of the former urban population is displaced to the rural area. The counterfactual income in the rural area is then again equal to the real income in the marginal city with the lowest real income (note: the marginal city with the lowest real income is the same with and without fiscal transfers). Hence, the spatial equilibrium condition is retained again, i.e., $\hat{v}_{rt} = \underline{\hat{v}}_{it}$ with $\hat{v}_{rt} = A_{rt}{\widehat{N}_{rt}}^{-\lambda}$.

Real income changes. For the change in real income, it is assumed that all residents that migrate to a new city become newcomers in these cities in the counterfactual scenario. The real income change for the rural population and newcomers (independently of whether they remain at their location or move to a new city) is given by $\hat{v}_{rt}/v_{rt}^* - 1$. The real income change for remaining incumbents in cities is given by $\hat{v}_{it}/v_{it}^* - 1$. The real income change for incumbents displaced from city j is given by $\hat{v}_{rt}/v_{jt}^* - 1$ independently of whether they move to a former donor city and become newcomers there or move to the rural area.