Cities with forking paths? Agglomeration economies in New Zealand 1976-2018*

Stuart Donovan[1,2,3], Thomas de Graaff[1,3], Arthur Grimes[2,4], Henri L.F. de Groot[1,3], and David C. Maré[2]

[1]Department of Spatial Economics, Vrije Universiteit Amsterdam, The Netherlands
[2]Motu Economic and Public Policy Research, Wellington, New Zealand
[3]Tinbergen Institute, Amsterdam, The Netherlands
[4]Victoria University, Wellington, New Zealand

10th February 2022

Abstract

We consider whether external urban economic advantages (agglomeration economies) vary with time and space using detailed micro-data on 134 locations in New Zealand for the period 1976–2018. We find subtle temporal variation, with estimates of agglomeration economies peaking in 1991 and then falling by approximately 1 percentage point in the subsequent 15-years. Since 2006, however, estimates have remained broadly stable; the world has not been getting "flatter". Our results reveal more significant spatial variation: Large cities offer net benefits in production but not consumption, whereas small locations close to large cities ("satellites") experience agglomeration economies that are stronger than average.

Keywords: agglomeration economies, cities, productivity, consumption, New Zealand
JEL-classification: R11, R23, R30.

* The authors acknowledge financial support from "Building Better Homes Towns and Cities", National Science Challenge; extremely helpful comments from the Editor and two anonymous Referees; and data supplied by Kate Preston and Shaan Badenhorst. Stuart Donovan acknowledges support from Veitch Lister Consulting and advice from Peter Nunns. Access to the data used in this study was provided by Statistics New Zealand under conditions designed to give effect to the security and confidentiality provisions of the Statistics Act 1975. The results presented in this study are the work of the authors, not Statistics New Zealand nor their data suppliers. Corresponding author: s.b.donovan@vu.nl.

1 Introduction

In "The Garden of Forking Paths," Borges uses a parable to illustrate the multiplicity of possible outcomes that can arise from the passage of time, observing "... time forks perpetually toward innumerable futures" (Borges et al., 1962). Analogies of this theme have been invoked to highlight barriers to the replication of scientific research and selection biases affecting publication (Gelman and Loken, 2013; Kasy, 2021). In economics, where empirical research often seeks to identify the causal effects of parameters that are subsequently used to appraise hypothetical scenarios and inform the policy process, we suggest the "forking paths" zeitgeist raises related questions about validity. That is, to what extent are economic parameters valid outside of the context—specifically, time and place—in which they are estimated? Researchers have begun to grapple with these gnarly epistemological questions (see, e.g., Meier and Sprenger, 2015; Rosenzweig and Udry, 2020). In this spirit, we consider the temporal stability and spatial transferability of estimates of external urban economic advantages, or "agglomeration economies". The complex microeconomic channels through which agglomeration can influence the choices of firms and households hint at a large number of possible "forking paths".

Why might policy-makers and researchers be interested in the stability and transferability of agglomeration economies? We posit three main reasons. First, estimates of agglomeration economies are often used to appraise the effects of policies, such as transport projects, in hypothetical scenarios that extend decades into the future. In doing so, appraisal guidelines usually assume that agglomeration economies are both stable and transferable—at least after controlling for differences in industrial composition (see, e.g., Table 37 in Waka Kotahi, 2020). Second, trends in agglomeration economies are central to debates on the future of cities. Whereas Cairncross (1997) and Friedman (2006) argued that technology will reduce the benefits of proximity, Glaeser (2011) opined that "despite the technological breakthroughs that have caused the death of distance, it turns out that the world isn't flat; it's paved." In the wake of the COVID-19 pandemic, debates about cities—and whether their economic advantages will persist in the future—have surfaced anew (see, e.g., Shenker, 2020). Third, and as alluded to above, agglomeration economies are a rubric used to describe effects that arise via complex microeconomic channels, such as knowledge spillovers and labour market efficiencies (Puga, 2010). Understanding how agglomeration economies vary over time and space may help illuminate aspects of these underlying channels. For these reasons, we suggest that the stability and transferability of agglomeration economies is relevant to both policy and research.

Surprisingly few empirical studies, however, consider whether agglomeration economies vary systematically with time and space. In terms of temporal stability, Martínez-Galarraga et al. (2008) analyse agglomeration economies in production for Spain and find estimates decline in the period 1860–1999. Similarly, the recent meta-analysis by Donovan et al. (2021) draws on estimates from a large number of primary studies for the period 1960–2020 and finds evidence that agglomeration economies in production have fallen by 1–2 percentage points since 2000. More attention has been paid to spatial variation, with several meta-analyses—see, for example, Ahlfeldt and Pietrostefani (2019) and Donovan et al. (2021)—and primary studies—see, for example, Ahrend et al. (2017) and Chauvin et al. (2017)—allowing for and reporting differences between countries. Fewer studies, however, consider the potential for spatial variation *within* countries, which is likely to be more relevant to domestic policy settings. Notably for our purposes, Maré and Graham (2013) find some evidence that agglomeration economies in production vary between nine regions in New Zealand after controlling for the characteristics of individual firms.

Against this backdrop, we consider whether agglomeration economies in production and consumption vary with time and space using data on 134 locations in New Zealand for the period 1976–2018. Following Roback (1982), Gabriel and Rosenthal (2004), and Maré and Poot (2019), we specify a simple economic model that measures the relative advantages of locations to firms and workers in terms of prices, from which we can derive expressions for agglomeration economies in terms of elasticities. Although grounded in the existing literature, our study combines several useful attributes. Unlike Gabriel and Rosenthal (2004) and Martínez-Galarraga et al. (2008), but like Maré and Graham (2013) and Maré and Poot (2019), we use micro-data to control for the observed characteristics of individual workers and rented dwellings. Whereas Martínez-Galarraga et al. (2008) present estimates for just two recent periods, namely 1965–1979 and 1985–1999, our estimates cover nine intervals for the period 1976–2018. And, in contrast to the nine broad regions used in Maré and Graham (2013), we estimate agglomeration economies for 134 individual urban locations in New Zealand. We also show that our results are robust to commuting, in the spirit of Albouy and Lue (2015) and Ahlfeldt, Redding et al. (2015); endogeneity, like Combes, Duranton, Gobillon and Roux (2010) and Combes, Duranton and Gobillon (2019); and various sensitivity tests.

Three main insights emerge from our analysis. First, we confirm that micro-data yield smaller estimates of agglomeration economies in production, with wage elasticities that are approximately half those from aggregate data—a finding that aligns with the seminal work by Combes, Duranton and Gobillon (2008) and the recent meta-analysis by

Ahlfeldt and Pietrostefani (2019). More uniquely, we find that the gap between estimates derived from micro-data vis-á-vis aggregate data has not changed with time, implying sorting by workers is a fairly stable process—at least in New Zealand. Second, we find evidence of subtle temporal variation in estimates of agglomeration economies, which arises due to the effect of agglomeration on rent. Specifically, from a nadir in 1981, estimates peak in 1991, and then decline by approximately 1 percentage point—loosely corroborating the trend in Donovan et al. (2021).[1] Since 2006, however, our estimates of agglomeration economies remain broadly constant; contrary to popular claims, the world has not been getting "flatter". Finally, we find evidence that agglomeration economies vary systematically between locations: Large cities in New Zealand offer net benefits in production but not in consumption, whereas "satellite" locations that are close to large cities experience agglomeration economies that are stronger than average. Evidence of large spatial differences within a small country like New Zealand—after controlling for the observed characteristics of workers and dwellings—is a notable finding. That said, agglomeration economies also exhibit a form of convergence, or regression to the mean, with estimates approaching the average as agglomeration increases.

Our findings highlight several enticing avenues for further research. First, we see value in primary research, like Martínez-Galarraga et al. (2008) and the present study, which uses consistent data and methods to trace the evolution of agglomeration economies within countries and over time. Second, further research would ideally seek to explain temporal and spatial variation in estimates of agglomeration economies. Although we detect the fingerprints of such variation, and present informal explanations for its origins, we do not pin-point the underlying causes. Third, we suggest there is merit in developing more sophisticated measures of agglomeration that, for example, consider multi-modal transport costs, rather than just car travel-time or distance. In New Zealand, for example, the combination of international remoteness, complex geography, and lack of inter-city passenger rail services may increase the relative importance of access to airports (Chen et al., 2021). Finally, we echo others by calling for more research into agglomeration economies in consumption, which appear more tenuous than those in production.

The remainder of this paper reads as follows: Section 2 develops the methodology; Section 3 summarises the data; Section 4 presents results, including sensitivity tests; Section 5 discusses our findings; and Section 6 concludes.

[1] In addition to the general factors discussed in Donovan et al. (2021), such as information and communications technologies ("ICT"), we speculate that temporal variation in New Zealand in the period 1976–2001 may reflect the lingering effects of major economic policy reforms that were implemented in the 1980s.

2 Methodology

2.1 Model

2.1.1 Basic model

Following Maré and Poot (2019) and Gabriel and Rosenthal (2004), the Basic model assumes mobile workers choose their home location, i, in period, t, given their preferences, U_{it}, for housing, H_{it}, a composite good, Y_{it}, and local amenities, $f_u(A_{it})$:

$$U_{it} = f_u(A_{it}) H_{it}^{\alpha} Y_{it}^{1-\alpha}. \tag{1}$$

Full-time workers supply one unit of labour and earn a wage, w_{it}. Whereas the price of housing, r_{it}, is locally determined, the price of the composite good is the same everywhere—so it is indexed in period t without loss of generality. From Equation (1) and the budget constraint, $w_{it} = r_{it} H_{it} + Y_{it}$, we derive the usual demand functions:

$$H_{it}^{*} = \alpha \frac{w_{it}}{r_{it}} \quad \text{and} \quad Y_{it}^{*} = (1-\alpha) w_{it}. \tag{2}$$

Substituting H_{it}^{*} and Y_{it}^{*} from Equation (2) into Equation (1) and re-arranging yields:

$$v_{it} = \kappa_u f_u(A_{it}) \frac{w_{it}}{r_{it}^{\alpha}} = \overline{v}_t, \tag{3}$$

where v_{it} denotes indirect utility and $\kappa_u = \alpha^{\alpha}(1-\alpha)^{1-\alpha}$ is a constant. Following Roback (1982), we impose a spatial equilibrium condition such that workers in all locations achieve the same fixed reservation utility, $\overline{v}_t$, in each period.

In terms of production, we assume that there is a representative firm in each location that uses a commonly-available constant returns to scale technology to produce Y_{it} using floorspace, H_{it}; labour, L_{it}; and mobile capital, K_{it} as inputs. Formally, we have:

$$Y_{it} = f_y(A_{it}) H_{it}^{\gamma_1} L_{it}^{\gamma_2} K_{it}^{1-\gamma_1-\gamma_2}, \tag{4}$$

where $f_y(A_{it})$ describes the effect of local amenities, A_{it}, on productivity. We assume firms pay k_t, r_{it}, and w_{it} for capital, floorspace, and labour, respectively.[2] Assuming firms maximise profits, we can derive the following equilibrium condition:

$$r_{it}^{\gamma_1} w_{it}^{\gamma_2} k_t^{1-\gamma_1-\gamma_2} = \kappa_y f_y(A_{it}), \tag{5}$$

where $\kappa_y = \gamma_1^{\gamma_1}\gamma_2^{\gamma_2}(1-\gamma_1-\gamma_2)^{1-\gamma_1-\gamma_2}$. Taking logs and re-arranging Equations (3) and (5) yields iso-utility and iso-cost conditions for workers and firms:

$$\ln f_u(A_{it}) = \alpha \ln r_{it} - \ln w_{it} + \ln \overline{v}_t - \ln \kappa_u, \tag{6}$$

$$\ln f_y(A_{it}) = \gamma_1 \ln r_{it} + \gamma_2 \ln w_{it} + (1-\gamma_1-\gamma_2)\ln k_t - \ln \kappa_y. \tag{7}$$

Equations (6) and (7) measure the implicit value of amenities to workers and firms in equilibrium as functions of structural parameters and the prices of inputs into utility and production. To understand the effect of agglomeration, $\ln D_{it}$, on the implicit value of amenities to workers and firms, we differentiate Equations (6) and (7) to arrive at:

$$\frac{\partial \ln f_u(A_{it})}{\partial \ln D_{it}} = \alpha \frac{\partial \ln r_{it}}{\partial \ln D_{it}} - \frac{\partial \ln w_{it}}{\partial \ln D_{it}} = \alpha \epsilon_{it}^r - \epsilon_{it}^w = E_c, \tag{8}$$

$$\frac{\partial \ln f_y(A_{it})}{\partial \ln D_{it}} = \gamma_1 \frac{\partial \ln r_{it}}{\partial \ln D_{it}} + \gamma_2 \frac{\partial \ln w_{it}}{\partial \ln D_{it}} = \gamma_1 \epsilon_{it}^r + \gamma_2 \epsilon_{it}^w = E_p, \tag{9}$$

where the space-invariant terms $\ln \overline{v}_t$, $\ln k_t$, $\ln \kappa_u$, and $\ln \kappa_y$ drop out and we let $\frac{\partial \ln r_{it}}{\partial \ln D_{it}} = \epsilon_{it}^r$ and $\frac{\partial \ln w_{it}}{\partial \ln D_{it}} = \epsilon_{it}^w$, that is, elasticities of rents and wages with respect to agglomeration, $\ln D_{it}$, which—in their most general form—can vary with location, i, and time, t. Together, Equations (8) and (9) conveniently express the change in the implicit value of local amenities that follows from a marginal change in $\ln D_{it}$ in terms of the latter's effects on rents and wages, as measured by their elasticities.[3] One direct implication of Equation (8) is that we can expect positive agglomeration economies in consumption, E_c, if and only if $\alpha \epsilon_{it}^r > \epsilon_{it}^w$. Bringing the Basic model to data then involves a relatively simple two-step process: First, we estimate the elasticities of rents and wages with respect to $\ln D_{it}$, ϵ_{it}^r and ϵ_{it}^w, and, second, we combine these elasticities as per Equations (8) and (9), given values for the structural parameters α, γ_1, and γ_2.

[2] Using r_{it} as the price of floorspace for firms may cause bias if the mix of inputs into floorspace differs between workers and firms and the prices of inputs varies over time. We expect this bias to be small.

[3] Specifically, Equation (8) defines the change in implicit value for workers who live and work in location, i.

2.1.2 Commuting model

We follow Albouy and Lue (2015) and introduce commuting into the Basic model. To begin, we assume that workers choose their home location based on the gross wages at the place-of-residence, w_{it}^h, net of average commuting costs, $\hat{C}_{it} \geq 1$, such that net wages are given by $w_{it}^h/\hat{C}_{it}$.[4] The analogous iso-utility condition to Equation (6) is then:

$$\ln f_u(A_{it}) = \alpha \ln r_{it} - \ln w_{it}^h + \ln \hat{C}_{it} + \ln \overline{v}_t - \ln \kappa_u, \tag{10}$$

Differentiating Equation (10) with respect to agglomeration yields $\frac{\partial \ln f_u(A_{it})}{\partial \ln D_{it}} = \alpha \epsilon_{it}^r - \epsilon_{it}^{w^h} + \epsilon_{it}^c = E_c$. This expression for E_c is similar to Equation (8) but includes the elasticity of wages at the place-of-residence, $\epsilon_{it}^{w^h}$, and the elasticity of average commute costs, ϵ_{it}^c.

Unlike wages and rents, we do not observe average commuting costs, $\ln \hat{C}_{it}$, in our data.[5] Instead, we proceed to construct $\ln \hat{C}_{it}$ from the following approximations:

$$w_{it}^h \approx \sum_j x_{jt|i}\, w_{jt} \qquad \text{and} \qquad \hat{C}_{it} \approx \sum_j x_{jt|i}\, C_{ij}. \tag{11}$$

In Equation (11), $x_{jt|i}$ denotes the share of workers that reside in i and work in j in period t. The left-hand approximation relates w_{it}^h to $x_{jt|i}$ and w_{jt}, whereas the right-hand approximation relates $\hat{C}_{it}$ to $x_{jt|i}$ and the commute costs from i to j, C_{ij}.[6] For observed values of w_{it}^h and w_{jt}, we can use the left-hand approximation to estimate $x_{jt|i}$, which can then be used in the right-hand approximation to construct $\hat{C}_{it}$.

We allow commuting shares, $x_{jt|i}$, to be endogenously determined with the wages available to workers, w_{jt}, and commuting costs, C_{ij}. Formally, we model $x_{jt|i}$ using a simple random utility sub-model of work location choice. This sub-model assumes the indirect utility, $v_{jt|i}$, that workers derive from working in location j in period t—conditional on residing in i—is given by $v_{jt|i} = \frac{w_{jt}}{C_{ij}} \varepsilon_j$, where ε_j denotes workers' unobserved preference for work location j. In spatial equilibrium, the expected value of $v_{jt|i}$ is the same for all

[4] That is, we assume that commuting costs are paid by workers, not firms, which leaves the production side of the model unchanged. In Section 4.2, we test whether there is empirical support for this assumption.
[5] Unfortunately, our data does not contain information on the commutes made by workers.
[6] Equation (11) are approximations because they exclude the effects of the rural "reservation". Given the low wages and long distances associated with working in rural areas, we expect that these commutes represent a relatively small share of workers—such that the error introduced by their omission is small.

alternative work locations, that is, $\mathbb{E}\left(\frac{w_{it}}{C_{ii}}\right) = \mathbb{E}\left(\frac{w_{jt}}{C_{ij}}\right) = \ldots$. Following Fosgerau and Bierlaire (2009), we assume ε_j are i.i.d. and observe an EV1 distribution, which allows us to derive the following expression for commuting shares

$$x_{jt|i} = \frac{\left(\frac{w_{jt}}{C_{ij}}\right)^{\lambda}}{\sum_k \left(\frac{w_{kt}}{C_{ik}}\right)^{\lambda}}, \tag{12}$$

where λ denotes a scale parameter to be estimated. Substituting the expression for $x_{jt|i}$ from Equation (12) into the two approximations in Equation (11) then yields:

$$w_{it}^{h} = \sum_j \frac{\left(\frac{w_{jt}}{\exp \beta t_{ij}}\right)^{\lambda}}{\sum_k \left(\frac{w_{kt}}{\exp \beta t_{ik}}\right)^{\lambda}} w_{jt} \qquad \text{and} \qquad \hat{C}_{it} = \sum_j \frac{\left(\frac{w_{jt}}{\exp \beta t_{ij}}\right)^{\lambda}}{\sum_k \left(\frac{w_{kt}}{\exp \beta t_{ik}}\right)^{\lambda}} \exp \beta t_{ij}. \tag{13}$$

In Equation (13), we let $C_{ij} = \exp \beta t_{ij}$ where t_{ij} denotes car travel-time between i and j, as per Ahlfeldt, Redding et al. (2015). As we observe w_{it}^{h}, w_{jt}, and t_{ij}, the left-hand side equation implicitly defines the parameters β and λ.[7] In this way, we can use observed wages at the place-of-residence and place-of-work, as well as data on travel-times, to construct a theoretically consistent measure of average commute costs, $\hat{C}_{it}$.

To finish the Commuting model, we relate the elasticity of wages at the place-of-residence, $\epsilon_{it}^{w^h}$, to that at places-of-work, ϵ_{jt}^{w}. Drawing on the above definitions, we observe:

$$\epsilon_{it}^{w^h} = \frac{\partial \ln w_{it}^{h}}{\partial \ln D_{it}} = \frac{\partial \ln \left(\sum_j x_{jt|i} w_{jt}\right)}{\partial \ln D_{it}}. \tag{14}$$

Equation (14) notes elasticities at the place-of-residence, ϵ^{w^h}, are a weighted average over elasticities at places-of-work, ϵ^{w}, where weights are defined by commuting shares, $x_{jt|i}$. Rather than deriving an analytical expression to map elasticities at the place-of-work to the place-of-residence, however, we instead use our data to estimate both elasticities—given that we observe both $\ln w_{it}^{h}$ and $\ln w_{it}$. Our approach to estimating the Commuting model is described in more detail in Section 2.2.2.

[7] Section 3.3 presents details of the grid search used to solve the left-hand side equation in Equation (13) for β and λ. In doing so, we find parameter estimates for β and λ that are similar to those reported in Ahlfeldt, Redding et al. (2015), with which our Commuting model shares some common elements.

2.2 Estimation

2.2.1 Basic model

To identify ϵ_{it}^{w} and ϵ_{it}^{r} in the Basic model, we estimate equations of the following form:

$$\ln w_{it} = f^{w}(\ln D_{it}; \epsilon_{it}^{w}) + \tau_{t}^{w} + \zeta_{i}^{w}$$
$$\ln r_{it} = f^{r}(\ln D_{it}; \epsilon_{it}^{r}) + \tau_{t}^{r} + \zeta_{i}^{r},$$

where $f(\ln D_{it}; \epsilon_{it})$ denotes a function of agglomeration, $\ln D_{it}$, which can vary between equations and models; τ_t denotes individual time effects; and ζ_i denotes individual location effects. For $\ln w_{it}$ and $\ln r_{it}$, we begin by using average income and rent. In our preferred models, however, we replace these averages with estimates of the spatial income and rent premia, which are derived from a process similar to that used in Combes, Duranton and Gobillon (2008).[8] Specifically, we first regress the reported incomes and rents of individual workers and dwellings for each Census year versus a set of location fixed effects and controls for observed individual characteristics. We estimate separate regressions for each year, such that the effects of observed characteristics on incomes and rents can vary with time. The location fixed effects—or, "spatial income and rent premia"—then become the dependent variables of the equations in our models.

We use Bayesian methods to estimate all models, which addresses four inter-related empirical problems.[9] First, we follow the literature on partial pooling models and treat individual time and location effects, τ_t and ζ_i, as group effects (or, "random effects") that are drawn from common population distributions with their own variances, or hyper-parameters (Gelman, Carlin et al., 2013). Compared to population effects (or, "fixed effects"), group effects pool information both within and between time periods and locations, mitigating the risk of over-fitting. With at most nine observations per location, we are especially concerned about over-fitting the location effects, ζ_i. And, unlike maximum likelihood estimators, Bayesian methods directly estimate the hyper-

[8] Unlike Combes, Duranton and Gobillon (2008), our data do not follow workers or dwellings over time, preventing us from including individual effects in the first-stage regressions. Donovan et al. (2021) find that including such effects reduces estimates of agglomeration economies in production by 0.011 whereas Ahlfeldt and Pietrostefani (2019) conclude that controlling for "sorting" reduces estimates by half.

[9] Specifically, all models are estimated using the statistical package `R` running in the RStudio environment with the `brms` package and default priors (R Core Team, 2021; RStudio Team, 2021; Bürkner, 2017).

parameters for the group effects. Second, our dependent variables—that is, the spatial income and rent premia—are measured with uncertainty, or error. Using Bayesian methods, specifically errors-in-outcomes models, we can account for uncertainty in the spatial premia—as measured by their estimated standard errors (s.e.)—in a direct and theoretically consistent manner.[10] Third, inspection of our data reveals the presence of some unusual observations that are associated with small towns (c.f. Figures 1 and 2). To mitigate the influence of these observations, we allow our two response variables to follow Student's t-distributions, which—compared to Gaussian distributions—allow more mass in the tails of the probability distributions (Geweke, 1993; Gelman and Hill, 2007). Finally, Equations (8) and (9) define agglomeration economies as composite parameters that are formed from estimates of both ϵ^w_{it} and ϵ^r_{it}. By jointly estimating these equations using Bayesian methods, the posterior distributions of parameter estimates for ϵ^w_{it} and ϵ^r_{it} can directly account for correlations between the two equations.[11]

We estimate four benchmark models. Model A uses average incomes and rents and imposes common elasticities, ϵ, across all time periods and locations:

$$\begin{aligned} \ln w_{it} &\sim \mathcal{N}(\epsilon^w \ln D_{it} + \tau^w_t + \zeta^w_i, \sigma^w) \\ \ln r_{it} &\sim \mathcal{N}(\epsilon^r \ln D_{it} + \tau^r_t + \zeta^r_i, \sigma^r), \end{aligned} \tag{Model A}$$

where we make the usual assumption that group effects, τ and ζ, are drawn from independent normal distributions. Similarly, Model B imposes common elasticities but instead uses income and rent premia within a multi-level, errors-in-outcomes setting:

$$\begin{aligned} \ln w_{it} &\sim \mathcal{N}(\ln w^*_{it}, s^w_{it}) \\ \ln r_{it} &\sim \mathcal{N}(\ln r^*_{it}, s^r_{it}) \\ \ln w^*_{it} &\sim \mathcal{N}(\epsilon^w \ln D_{it} + \tau^w_t + \zeta^w_i, \sigma^w) \\ \ln r^*_{it} &\sim \mathcal{N}(\epsilon^r \ln D_{it} + \tau^r_t + \zeta^r_i, \sigma^r), \end{aligned} \tag{Model B}$$

where we assume the estimated spatial premia, $\ln w_{it}$ and $\ln r_{it}$, are drawn from normal distributions with true means, $\ln w^*_{it}$ and $\ln r^*_{it}$, and standard deviations, s^w_{it} and s^r_{it},

[10] Combes, Duranton and Gobillon (2008) address errors-in-outcomes using a feasible generalized least-squares ("FGLS") estimator, which yields similar results to OLS estimators. The authors attribute this result to the large number of observations at their disposal. Subsequent researchers have often followed this lead, pointing to precise estimates of the spatial premia as reason to treat the latter deterministically (see, e.g., footnote 13 in Groot et al., 2014). In our case, however, the relatively large number of small locations in our data often leads to somewhat imprecisely estimated spatial premia.

[11] Joint estimation, for example, allows us to model correlations between the residuals and individual effects, τ_t and ζ_i. In Section 4.3.3, we model these correlations as a sensitivity test.

respectively, as defined by the standard errors from the first-stage regressions.[12]

Model C then allows the true means, $\ln w_{it}^*$ and $\ln r_{it}^*$, to follow Student's t-distributions with additional degrees-of-freedom parameters, ν:

$$\begin{aligned}
\ln w_{it} &\sim \mathcal{N}(\ln w_{it}^*, s_{it}^w) \\
\ln r_{it} &\sim \mathcal{N}(\ln r_{it}^*, s_{it}^r) \\
\ln w_{it}^* &\sim t(\epsilon^w \ln D_{it} + \tau_t^w + \zeta_i^w, \nu^w, \sigma^w) \\
\ln r_{it}^* &\sim t(\epsilon^r \ln D_{it} + \tau_t^r + \zeta_i^r, \nu^r, \sigma^r).
\end{aligned} \tag{Model C}$$

Finally, Model D allows elasticities to vary with time and location, as per the group effects, δ_t and γ_i, respectively. In its most general form, Model D is defined as

$$\begin{aligned}
\ln w_{it} &\sim \mathcal{N}(\ln w_{it}^*, s_{it}^w) \\
\ln r_{it} &\sim \mathcal{N}(\ln r_{it}^*, s_{it}^r) \\
\ln w_{it}^* &\sim t((\epsilon^w + \delta_t^w + \gamma_i^w) \ln D_{it} + \tau_t^w + \zeta_i^w, \nu^w, \sigma^w) \\
\ln r_{it}^* &\sim t((\epsilon^r + \delta_t^r + \gamma_i^r) \ln D_{it} + \tau_t^r + \zeta_i^r, \nu^r, \sigma^r).
\end{aligned} \tag{Model D}$$

In Section 4, we estimate several variants of Model D that restrict $\delta_t = 0$ and $\gamma_i = 0$.

In short, Model A and Model B differ only in the use of aggregate data vis-á-vis spatial premia, where the latter are preferred but require that we model errors-in-outcomes. Model C then mitigates the effects of influential observations by allowing the dependent variables to follow Student's t-distributions. Despite their differences, Model A, Model B, and Model C all assume that income and rent elasticities are constant with time and location. In contrast, Model D allows elasticities to vary with the time and location group effects, δ_t and γ_i. We are primarily interested in understanding whether there is evidence to prefer variants of Model D, where the income and rent elasticities can vary with time and location, over Model C, where they cannot. And, in keeping with our aforementioned interest in the external validity of estimates of agglomeration economies, we focus on the out-of-sample predictive performance of these benchmark models, that is, their ability to predict observations not used in their estimation.

[12] That is, $\ln w_{it}$ and $\ln r_{it}$ are realisations of random variables with underlying distributions, for which the latent means, $\ln w_{it}^*$ and $\ln r_{it}^*$, are estimated based on the information in the wider linear model. In this way, "statistical shrinkage" reduces the influence of unusual observations in a consistent manner.

2.2.2 Commuting model

We estimate the Commuting model developed in Section 2.1.2 in a similar way to that for the Basic model above, but with two additional equations that allow us to estimate the elasticity of wages at the place-of-residence, $\epsilon_{it}^{w^h}$, and the elasticity of average commuting costs, ϵ_{it}^{c}. Specifically, we estimate models with the following general form:

$$
\begin{aligned}
\ln w_{it} &= f^{w}(\ln D_{it}, \ln \hat{C}_{it}; \epsilon_{it}^{w}) + \tau_t^{w} + \zeta_i^{w} \\
\ln w_{it}^{h} &= f^{w^h}(\ln D_{it}, \ln \hat{C}_{it}; \epsilon_{it}^{w^h}) + \tau_t^{w^h} + \zeta_i^{w^h} \\
\ln r_{it} &= f^{r}(\ln D_{it}, \ln \hat{C}_{it}; \epsilon_{it}^{r}) + \tau_t^{r} + \zeta_i^{r} \\
\ln \hat{C}_{it} &= f^{c}(\ln D_{it}; \epsilon_{it}^{c}) + \tau_t^{c} + \zeta_i^{c}
\end{aligned}
$$

In this general specification of the Commuting model, we include our estimate of average commuting costs, $\ln \hat{C}_{it}$, in the income and rent equations for three reasons. The first is that shocks to $\ln \hat{C}_{it}$ may be correlated with shocks to incomes and rents, such that omitting $\ln \hat{C}_{it}$ could introduce bias into our estimates of ϵ_{it}^{w}, $\epsilon_{it}^{w^h}$, and ϵ_{it}^{r}. The second reason is that including $\ln \hat{C}_{it}$ in the equations for incomes and rents allows us to check some of the theoretical assumptions of the model. Specifically, our model assumes $\ln \hat{C}_{it}$ affects workers but not firms. In practice, however, $\ln \hat{C}_{it}$ may affect production, for example because it acts as a proxy for congestion effects or because it undermines the efficient functioning of labour markets.[13] The third reason we include $\ln \hat{C}_{it}$ in the equation for rents is that it aligns our model more closely with Albouy and Lue (2015), where commuting costs are found to have a negative effect on rents.

We are primarily interested in whether the Basic and Commuting models yield similar predictions for spatial and temporal variation in estimates of agglomeration economies in production and consumption, E_p and E_c. Put another way, we seek to understand whether both models imply similar spatial and temporal patterns, if any. We consider this question in Section 4.2, where we compare the predictions of the Commuting model to those of the preferred Basic model—both on average and for individual locations.

[13] Labour market imperfections could reflect market power held by either firms or workers. For example, where firms have monopsony power, we might expect average commuting costs to negatively affect $\ln w_{it}$ and $\ln w_{it}^{h}$, as firms bid down wages in response to workers having fewer outside employment options. Alternatively, where workers have bargaining power, we might expect average commute costs to positively affect $\ln w_{it}$ and $\ln w_{it}^{h}$, as workers pass through commuting costs to their employer.

3 Data

3.1 Income, rent, and population

We source income, rent, and population data from the New Zealand Census ("Census"). The cross-sectional dimension is defined by the 143 zones of the urban area classification developed by Statistics New Zealand, where we consolidate zones in metropolitan areas, namely Auckland (four zones), Wellington (four zones), Hamilton (three zones), and Napier-Hastings (two zones).[14] The temporal dimension is defined by the nine waves from 1976 to 2018.[15] For the 1,206 observations in our panel—that is, 134 locations over nine years—we extract data on all full-time workers aged 25-years or older and all rented dwellings. Due to security and confidentiality provisions, we lose 26 observations associated with small towns in early Censuses—leaving us with 1,180 observations.[16] We index incomes and rents to 2018 (second quarter) levels using consumer price indices. In 2018, the median and mean population of locations is 4,464 and 30,189, respectively.

Census data confer advantages and disadvantages. On the upside, the Census seeks to survey all residents and typically achieves response rates above 95%, mitigating problems with sample selection. And, by accessing unit records, we can control for individual characteristics at the place-of-residence and place-of-work, both of which feature in our model. The Census also has some limitations, however. As the Census does not track individuals over time, we cannot control for unobserved sources of heterogeneity. Data on income and rents are grouped into \$100 and \$10 bands, introducing some measurement error.[17] Unfortunately, the Census records gross income not labour income, when the latter is more relevant to our model.[18] Finally, the Census only collects prices for rented dwellings, which represent approximately one-third of dwellings over this period.

[14] Stats NZ (2017, p. 14) describes urban areas as having " ... high population density with many built environment features where people and buildings are located close together ... " We assign records to 2018 urban areas using the most detailed geographies in each Census. Where geographies from an earlier Census are associated with more than one urban area, we allocate records to the area with the largest share of the 2018 population. Generally, this is a "meshblock", which contain approximately 100 people. For the 1976 Census, meshblocks were derived from administrative codes. For people who were away from home on Census night, coding is available only at a more aggregate ("area unit") level.

[15] Specifically, the Census was undertaken in 1976, 1981, 1986, 1991, 1996, 2001, 2006, 2013, and 2018.

[16] Specifically, we lose 13, 10, and 3 observations from the 1976, 1981, and 1986 Censuses, respectively.

[17] We assign responses to the midpoint of bands and top-code responses in the highest band.

[18] Gross income appears to be an accurate measure of labour income for full-time workers aged 25-years or older. For these workers, Maré and Poot (2019) regress gross income from the 2001, 2006, and 2013 Censuses versus labour income from administrative data and find unitary coefficients and $R^2 > 0.97$.

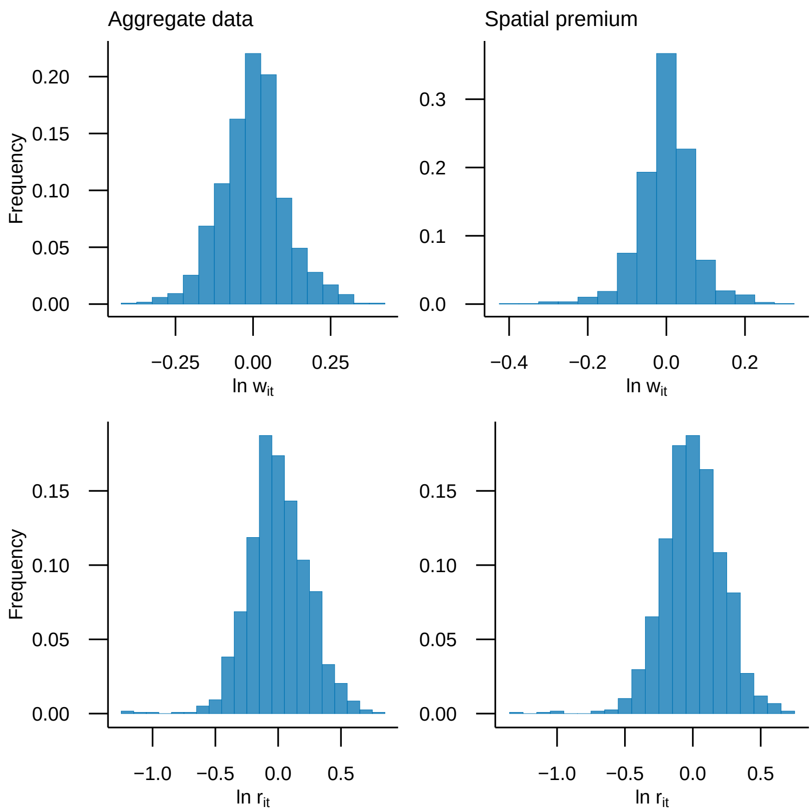

Figure 1: Histograms of dependent variables, mean-centred by year. Left and right panels show aggregate data and spatial premia; top and bottom panels show incomes and rents, respectively.

In Figure 1, the top and bottom panels show incomes and rents whereas the left and right panels show aggregate data and spatial premia, respectively. The spatial income premia control for gender, age (polynomial by gender), qualification ($n = 10$, except in 1976 when $n = 4$), two-digit industry sector ($n = 54$), ethnicity ($n = 15$), religion ($n = 11$), and birthplace ($n = 12$). The spatial rent premia control for the number of bedrooms ($n = 10$) and rooms ($n = 10$) as well as dwelling type ($n = 8$) and heating ($n = 8$). Regression results for the spatial income and rent premia are available on request.

3.2 Agglomeration

We follow Holl (2012) and measure agglomeration in terms of "effective population", $D_{it} = P_{it} + \sum_{j \neq i} \frac{P_{jt}}{d_{ij}}$, where the population, P_{it}, of location i is added to the sum of the population, P_{jt}, of other locations j inversely-weighted by the distance in kilometres

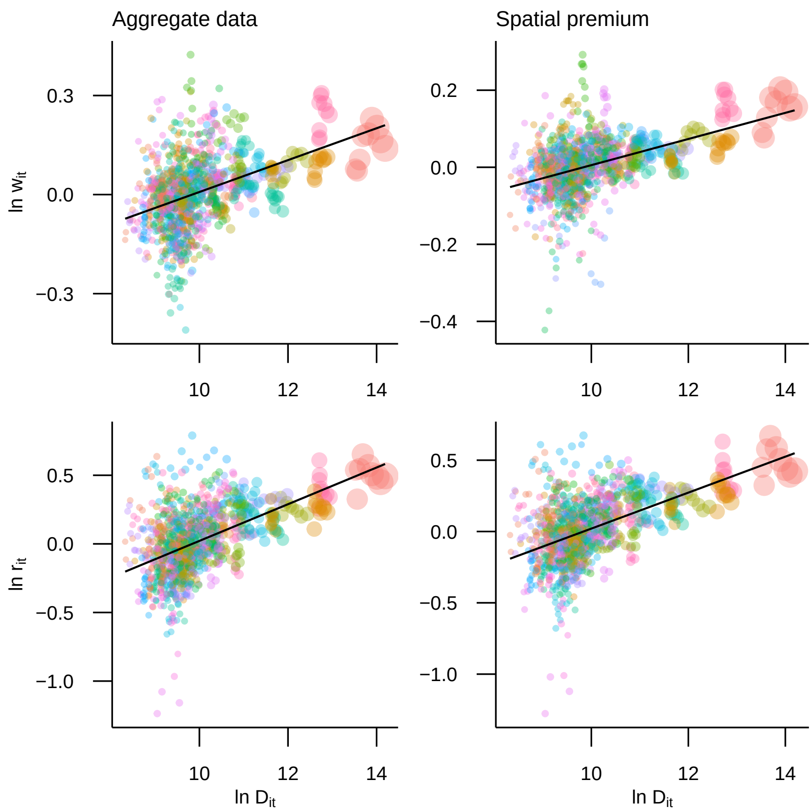

Figure 2: Scatter plots of $\ln w_{it}$ and $\ln r_{it}$ (vertical axes), mean-centred by year, versus $\ln D_{it}$ (horizontal axes), as defined in text. Left and right panels show aggregate data and spatial premia; top and bottom panels show incomes and rents, respectively. The size of points indicates population, P_{it}.

between their population-weighted centroids, d_{ij}. We are not aware of historical data on distances, d_{ij}, so use data for 2018. On average, approximately two-thirds of D_{it} is due to the own-population, P_{it}, with the balance due to other locations, P_{jt}. Figure 2 reveals the expected positive association between $\ln w_{it}$ and $\ln r_{it}$ (vertical axes) versus $\ln D_{it}$ (horizontal axes). Although we seek to estimate total agglomeration economies rather than distinguish between underlying microeconomic channels—such as knowledge spillovers, labour market efficiencies, and input-output (IO) linkages—the strength of these channels will define the spatial scope of effects. Evidence finds, for example, knowledge spillovers decay most rapidly (see, e.g., Autant-Bernard and LeSage, 2011), labour market efficiencies have a larger scope (see, e.g., Rice et al., 2006), and IO linkages even broader still (see, e.g., Ehrl, 2013). For firms, we expect $\ln D_{it}$ will capture Marshallian advantages arising via labour markets and IO linkages whereas, for workers, we expect $\ln D_{it}$ will capture the benefits of endogenous amenities (Gaigné et al., 2022; Gautier et al., 2010). Section 4.3.4 considers alternative measures of D_{it}.

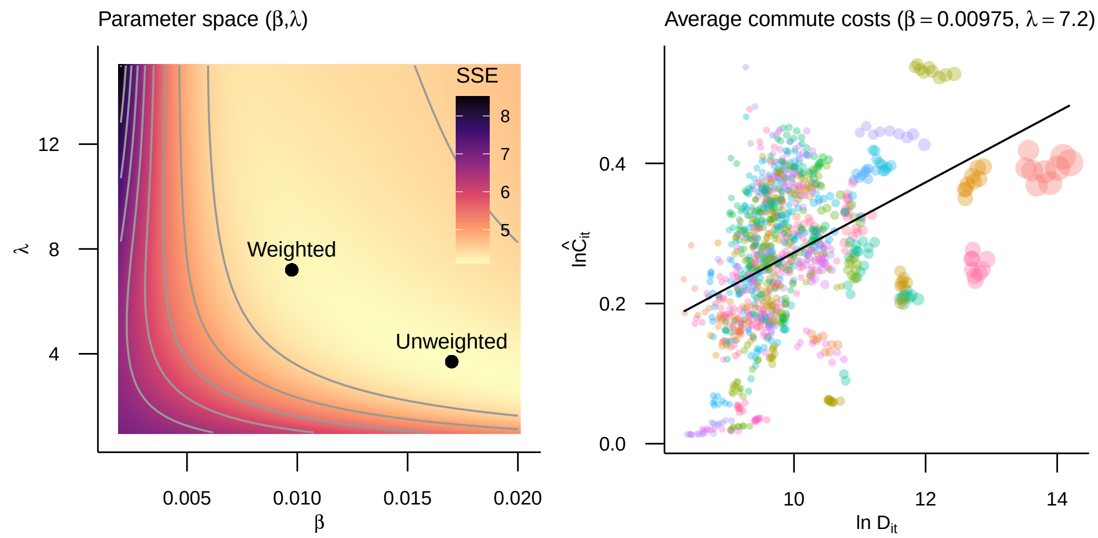

Figure 3: Left panel: (β, λ) parameter space, where the "Unweighted" point (0.017, 3.7) denotes that which minimises SSE, whereas the "Weighted" point (0.00975, 7.2) denotes that which minimises SSE when squared errors are weighted by the variance of $\ln w^h_{it}$. Right panel: Plots average commute costs $\ln \hat{C}_{it}$—constructed with the "Weighted" parameter values—versus agglomeration, $\ln D_{it}$.

3.3 Commuting costs

We construct average commuting costs, $\ln \hat{C}_{it}$, using the model in Section 2.1.2. Formally, we apply a grid search to the left-hand equation in Equation (13) to find the values of β and λ that minimise the sum of squared errors ("SSE") between observed and predicted values of w^h_{it}.[19] The left panel of Figure 3 shows the (β, λ) parameter space that results from this grid search whereas the right panel plots estimates of $\ln \hat{C}_{it}$ versus agglomeration, $\ln D_{it}$, for our preferred "weighted" parameter values (0.00975, 7.2).[20] As expected, we find a positive association between $\ln \hat{C}_{it}$ and $\ln D_{it}$, with considerable variation between locations. Using this calibration of the Commuting model, we estimate the share of the workers who live and work in the same location to be 78.5%. That is, the model predicts approximately three-quarters of workers commute internally. If accurate, then this implies that our data—specifically, the urban areas defined by Stats NZ (2017)—are suited to the Basic model, which focuses on home location choice.

[19] We estimate internal travel-times, t_{ii} by, first, calculating the average internal distance, $d_{ii} = \frac{2}{3}\sqrt{A_i/\pi}$, to the centre of a circle with area, A_i; second, scaling d_{ii} by a factor of 1.5 to approximate the distance by road; and, third, assuming average commuting speeds of 30 kilometres per hour.

[20] These values are almost identical to the (0.010, 6.8) reported in Ahlfeldt, Redding et al. (2015), with which our model shares some common elements. The SSE surface in the left panel of Figure 3, however, implies that a wide range of values for β and λ lead to similarly low values of SSE. In Section 4.2, we present results for $\ln \hat{C}_{it}$ constructed from both unweighted and weighted values of (β, λ).

4 Results

4.1 Basic model

Table 1 presents results for the four Basic models specified in Section 2.2. The first panel presents estimates of the common income and rent elasticities, ϵ^w and ϵ^r, whereas the second panel presents estimates of agglomeration economies, E_p and E_c, which we compute from ϵ^w and ϵ^r using Equations 8 and 9. For the structural parameters, we use our data to estimate the cost share of housing, $\alpha = 0.245$.[21] On the production side, we follow Maré and Poot (2019) and Fabling and Maré (2019) and assume $\gamma_1 = 0.10$ and $\gamma_2 = 0.63$ for the cost shares of floorspace and labour in production, respectively.[22] We note the assumptions for γ and α do not affect ϵ^w and ϵ^r but instead shift the levels of the estimates for agglomeration economies in production and consumption, E_p and E_c.[23] The third panel summarises the time- and location-specific elasticities that are included in variants of Model D. And, finally, the last panel summarises two model performance measures, R^2 and elpd, where we prefer the latter.[24]

Model A—which uses aggregate data—returns an income elasticity, ϵ^w, that is approximately twice as large as that for Model B—which uses micro-data, although the rent elasticity, ϵ^r, is largely unchanged.[25] Compared to Model A, Model B returns estimates of agglomeration economies in production, E_p, which are approximately 2.3 percentage points smaller. Model B performs worse than Model C, supporting the latter's use of Student's t-distributions to mitigate influential observations.[26] As for Model D, we find

[21] We estimate α by regressing aggregate data on $\ln r_{it}$ versus $\ln w_{it}$ with individual year and location group-effects. The resulting estimate for $\alpha = 0.245$ (s.e. 0.013) lies between the 0.20 and 0.30 values reported in MBIE (2015, p. 10) and Glaeser (2008), respectively.

[22] Fabling and Maré (2019) report aggregate Cobb Douglas production function estimates, where capital accounts for 14% of expenditure, labour 24% and intermediate purchases and taxes 62%. From this we can impute the labour share of factor payments, $\gamma_2 = 0.24/(0.14 + 0.24) = 0.63$.

[23] Specifically, $\frac{\partial E_p}{\partial \gamma} = \epsilon^r_{it} - \epsilon^w_{it}$ and $\frac{\partial E_c}{\partial \alpha} = \epsilon^r_{it}$. For example, if cost shares are 20% larger than we assume, such that $\gamma = 0.12$ and $\alpha = 0.288$, then the estimated E_p and E_c for Model D4 will shift upwards by $0.02(0.301 - 0.036) = 0.005$ and $0.048(0.301) = 0.015$, respectively.

[24] The Bayesian R^2 is the median value from the posterior predictions of a model, averaged across the two equations. The expected log pointwise predictive density ("elpd") measures the out-of-sample performance of a model using efficient leave-one-out cross-validation; see Vehtari et al. (2017) for details. We note the R^2 and elpd is only comparable between Model B, Model C, and Model D, which use the same dependent variables—that is, the spatial income and rent premia that are estimated in the first-stage regressions.

[25] We posit micro-data matters more for ϵ^w than ϵ^r because, first, workers are more mobile than dwellings and, second, we have weaker quality adjustments for dwellings.

[26] We find $\nu^w \approx 66$ and $\nu^r \approx 6$. The latter indicates considerable mass exists in the tails of the probability distribution for $\ln w_{it}$, providing further empirical support for Model C vis-á-vis Model B.

	Model A	Model B	Model C	Model D				
				D1	D2	D3	D4	D5
ϵ^w	0.071	0.037	0.036	0.036	0.036	0.036	0.036	0.036
	(0.009)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)
ϵ^r	0.182	0.165	0.182	0.191	0.384	0.181	0.301	0.303
	(0.019)	(0.017)	(0.019)	(0.024)	(0.055)	(0.018)	(0.054)	(0.052)
E_p	0.063	0.040	0.041	0.042	0.061	0.041	0.053	0.053
	(0.006)	(0.004)	(0.004)	(0.005)	(0.007)	(0.004)	(0.006)	(0.007)
E_c	-0.026	0.004	0.008	0.011	0.058	0.008	0.037	0.038
	(0.010)	(0.007)	(0.007)	(0.009)	(0.015)	(0.008)	(0.014)	(0.014)
δ_t^w	No	No	No	Yes	No	Yes	No	Yes
γ_i^w	No	No	No	No	Yes	Yes	No	Yes
δ_t^r	No	No	No	Yes	No	No	Yes	Yes
γ_i^r	No	No	No	No	Yes	No	Yes	Yes
R^2 [%]	84.9	99.2	99.2	99.3	99.4	99.2	99.4	99.4
elpd	2,154	1,914	1,970	2,012	2,188	1,970	2,214	2,213

Table 1: Estimates for the four Basic models (s.e. in parentheses). All models use 1,180 observations and include individual time and location effects, τ_t and ζ_i.

Model D4—in which ϵ^r varies with time and location—has the highest elpd, closely followed by Model D5. We find stable estimates for ϵ^w of 0.036–0.037 in Model B, Model C, and the five variants of Model D. For ϵ^r, we find somewhat larger estimates in Models D2, D4, and D5 than Model A, Model B, and Model C. Nonetheless, all estimates of ϵ^r lie close to the 0.176–0.304 range reported in Combes, Duranton and Gobillon (2019, c.f. Table 4, p. 1573). Variation in ϵ^r has economically meaningful implications: Whereas Model C predicts agglomeration economies in consumption, E_c, are approximately zero, Models D2, D4, and D5 predict they are positive. On balance, the results in Table 1 provide evidence for positive agglomeration economies in production, E_p, although more uncertainty exists for those in consumption, E_c. We take Model D4 forward for further testing, in which temporal and spatial variation arises via ϵ^r.[27]

4.2 Commuting model

In Table 2, we test the sensitivity of results to average commuting costs, $\ln \hat{C}_{it}$. The first four columns add $\ln \hat{C}_{it}$ to Basic model D4. Following Equation (13) and Figure 3, Variants 1 and 2 construct $\ln \hat{C}_{it}$ using "Unweighted" parameters whereas Variants 3 and 4 use "Weighted" parameters. Variants 1 and 3 add $\ln \hat{C}_{it}$ to the rent equation, r,

[27] We also estimate two variants of Model D4 where time- and location-specific effects, γ_i^w and δ_t^w, enter separately, which return similar results to Model D4. Results for these variants are available on request.

	Basic model D4 with commute costs				Commuting model			
	1	2	3	4	1	2	3	4
ϵ^w	0.036	0.048	0.036	0.049	0.037	0.049	0.036	0.049
	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
ϵ^r	0.328	0.328	0.309	0.309	0.325	0.328	0.311	0.307
	(0.051)	(0.054)	(0.052)	(0.053)	(0.053)	(0.052)	(0.052)	(0.052)
ϵ^{w^h}					0.035	0.039	0.035	0.041
					(0.005)	(0.005)	(0.005)	(0.005)
ϵ^c					0.013	0.013	0.008	0.008
					(0.009)	(0.009)	(0.006)	(0.007)
E_p	0.056	0.063	0.054	0.062	0.065	0.076	0.064	0.075
	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)
E_c	0.044	0.032	0.040	0.026	0.057	0.054	0.049	0.043
	(0.014)	(0.015)	(0.014)	(0.014)	(0.017)	(0.017)	(0.015)	(0.015)
$\ln \hat{C}$	r	r, w	r	r, w	r	r, w	r	r, w
(β, λ)	Un.	Un.	Wei.	Wei.	Un.	Un.	Wei.	Wei.
Rsq.	99.4	99.4	99.4	99.4	98.5	98.5	96.2	96.3
elpd	$2,221$	$2,241$	$2,219$	$2,260$	$6,313$	$6,335$	$6,434$	$6,483$

Table 2: Estimates for Basic and Commuting models (s.e. in parentheses). All models use 1,180 observations and include individual year and location effects. Variants 1 and 3 add $\ln \hat{C}$ to the rent equation whereas Variants 2 and 4 add $\ln \hat{C}$ to both the rent and wage equation. Variants 1 and 2 use unweighted β and λ parameters to construct $\ln \hat{C}$ whereas Variants 3 and 4 use weighted parameters.

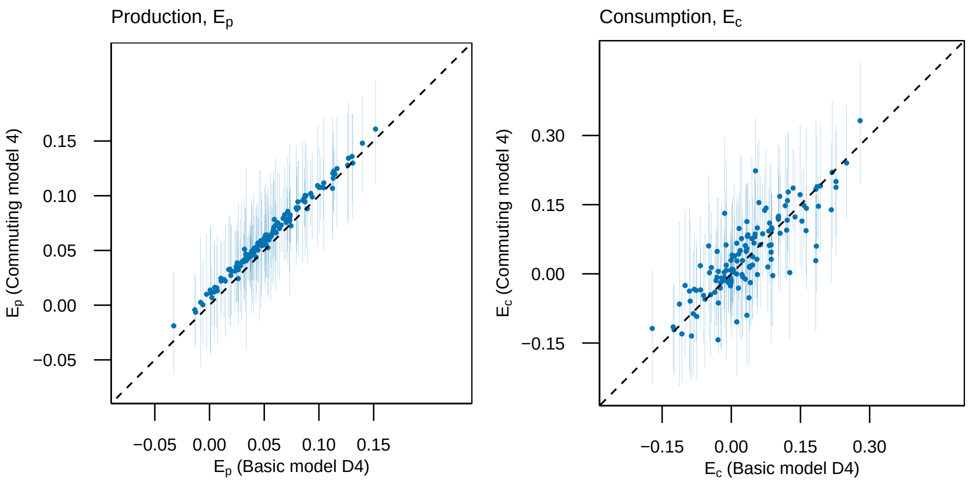

Figure 4: Location-specific estimates of E_p (left panel) and E_c (right panel) for Basic model D4 (horizontal axis) and Commuting model 4 (vertical axis). The dashed 45 degree line denotes the points where both models yield the same estimates and the error bars denote 95% credible intervals.

whereas Variants 2 and 4 add $\ln \hat{C}_{it}$ to both equations, r and w. The last four columns of Table 2 replicates these four variants for the Commuting model, which also estimates the elasticity of income at the place-of-residence, ϵ^{w^h}, and the elasticity of average commuting costs, ϵ^c. We find stable estimates for the income and rent elasticities as

	Model D	Sensitivity tests						
	4	S1	S2	S3	S4	S5	S6	S7
ϵ^w	0.036	0.031	0.036	0.037	0.037	0.038	0.025	0.025
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.004)	(0.003)
ϵ^r	0.301	0.269	0.289	0.299	0.326	0.298	0.215	0.172
	(0.054)	(0.046)	(0.049)	(0.055)	(0.041)	(0.055)	(0.031)	(0.024)
η			0.149					
			(0.030)					
E_p	0.053	0.046	0.052	0.053	0.056	0.054	0.038	0.033
	(0.006)	(0.006)	(0.006)	(0.007)	(0.005)	(0.007)	(0.004)	(0.003)
E_c	0.037	0.035	0.034	0.036	0.043	0.035	0.028	0.017
	(0.014)	(0.013)	(0.013)	(0.015)	(0.012)	(0.015)	(0.009)	(0.007)
Rsq. [%]	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.4
elpd	2,214	2,254	2,210	3,036	2,148	2,216	2,098	2,173

Table 3: Estimated elasticities for sensitivity tests with standard errors in parentheses. All models include individual year and location effects and use 1,180 observations except for Model S3, which uses 1,059 observations. See Sections 4.3.1 (S1), 4.3.2 (S2), 4.3.3 (S3), and 4.3.4 (S4–S7) for details.

well as the associated agglomeration economies, E_p and E_c.[28] In Figure 4, we compare location-specific estimates of E_p (left panel) and E_c (right panel) for Basic model D4 and Commuting model 4. We find similar estimates for E_p but larger differences between the estimates for E_c, which are attributable to the location-specific commute elasticities, ϵ_i^c. Few estimates of E_c, however, are significantly affected—as evident from the error bars.[29] Given the alignment between the predictions of the Basic model and those of the Commuting model, we choose to take the former forward for further sensitivity testing.

4.3 Sensitivity tests

4.3.1 Additional correlations

We extend Model D4 to allow for two additional types of correlation. First, we allow the error terms to be correlated between the income and rent equations, as in a seemingly

[28] We find $\ln \hat{C}_{it}$ has a negative effect on rents and incomes. The effect on rents aligns with Albouy and Lue (2015), although the effect on incomes is more interesting. We posit $\ln \hat{C}_{it}$ may act as a proxy for congestion or capture labour market imperfections, such as firms with monopsony power.

[29] Albouy and Lue (2015) find commute costs explain spatial variation in the *levels* of wages, rents, and commute costs. In contrast, we focus on spatial variation in the *marginal effects of agglomeration*, $\ln D$, on these economic outcomes. Levels of amenities may plausibly be affected by average commute costs, $\ln \hat{C}$, even if the latter is somewhat insensitive to changes in $\ln D$. Indeed, the right panel of Figure 3 reveals little variation in $\ln \hat{C}$ within locations over time, which may explain the small size of estimates for ϵ^c and the alignment in estimates of E_p and E_c between the Basic and Commuting models.

unrelated regression—or "SUR" (Zellner and Ando, 2010). Second, we allow for correlations between the time and location effects, τ_t and ζ_i, which enter both equations.[30] Results for this model ("S1") are shown in the second column of Table 3. We find a large, positive, and precisely-estimated correlation (0.76; s.e. 0.02) between the error terms, which suggests incomes and rents are jointly affected by unobserved shocks. We find a similarly large, positive, and precisely-estimated correlation between τ_t (0.69; s.e. 0.20), although no correlation between ζ_i. Compared to Model D4, we find Model S1 returns similar—albeit somewhat smaller—estimates for elasticities and agglomeration economies, E_p and E_c. The elpd of Model S1 (2,254) is higher than that for Model D4 (2,214), which suggests allowing for additional correlations between the two equations helps to improve model performance.

4.3.2 Income-elasticity of rents

We extend Model D4 to include the spatial income premium at the place-of-residence, $\ln w_{it}^h$, in the rent equation, which becomes $\ln r_{it}^* \sim \ t((\epsilon^r + \delta_t^r + \gamma_i^r) \ln D_{it} + \eta \ln w_{it}^{*h} + \tau_t^r + \zeta_i^r, \nu^r, \sigma^r)$.[31] This specification ("S2") allows us to estimate the income-elasticity of rents, η, and test whether housing expenditure is a constant share α of income, w_{it}—that is, η =1—as predicted by the Marshallian demand for housing in Equation 2, Section 2.1. Results for this model are shown in the third column of Table 3. We find similar elasticities to Model D4, although the estimated income elasticity of rents, η, has a median value of 0.15 and a 95% credible interval of [0.09, 0.21]. As the estimated value for η is less than one, this implies rents are not a constant share of income—specifically they are inelastic.[32] We note the elpd for Model S2 is lower than Model D4, which implies adding income to the rent equation brings about a deterioration in model performance.

[30] That is, we assume individual effects are drawn from bivariate normal distributions, $\tau \sim \mathcal{N}_2(0, \Sigma_\tau)$ and $\zeta \sim \mathcal{N}_2(0, \Sigma_\zeta)$, where off-diagonal elements of the covariance matrices, Σ_τ and Σ_ζ, can be non-negative.

[31] Uncertainty in $\ln w_{it}^h$ is modelled in a multi-level setting using the standard errors from the first-stage regressions, formally $\ln w_{it}^h \sim \ \mathcal{N}(\ln w_{it}^{*h}, s_{it}^{w^h})$. The latent variable, $\ln w_{it}^{*h}$, then enters the rent equation.

[32] Differences in the samples used to calculate the spatial income and rent premia may cause η to be biased downwards. Specifically, whereas income premia are derived from data on all full-time workers aged 25-years or older, rent premia pertain only to the subset of dwellings that are rented. As workers who own their own home are expected to have, on average, higher incomes than those who rent, locations with higher levels of home ownership are likely to have higher $\ln w_{it}^{*h}$ but not necessarily higher $\ln r_{it}^*$—causing η to be biased downwards. We do not expect this issue to affect the estimated rent elasticities, however.

4.3.3 Production of housing

We extend Model D4 to give an explicit role to the production of housing. To proceed, we assume the marginal cost of producing housing, m_{it}, is given by:

$$m_{it} = (P_{it})^{\rho_1} \left(\frac{P_{it}}{P_{it-1}} \right)^{\rho_2} (w_{it}^h)^{\rho_3} \rho_{4,i} \rho_{5,t}, \tag{15}$$

where P_{it} and P_{it-1} denote the population at time t and $t-1$, respectively; w_{it}^h denotes the income premia at the place-of-residence; $\rho_{4,i}$ and $\rho_{5,t}$ denote location and time effects; and ρ_1, ρ_2, and ρ_3 are parameters to be estimated. The elements of Equation (15) have intuitive interpretations. First, marginal costs, m_{it}, increase with population, P_{it}, as more cost-effective housing sites are developed first. Second, we allow for adjustment frictions that may cause m_{it} to increase with investment P_{it}/P_{it-1}, as in Glaeser et al. (2014). Third, like Section 4.3.2 we allow m_{it} to increase with income, w_{it}^h. Finally, $\rho_{4,i}$ and $\rho_{5,t}$ capture spatial and temporal factors that affect m_{it}, such as geography and policy.

Assuming perfectly competitive housing markets, such that $m_{it} = r_{it}$, taking logs of Equation (15), and re-arranging for $\ln P_{it}$, yields the following equilibrium condition:

$$\ln P_{it} = \frac{1}{\rho_1 + \rho_2} \left(\ln r_{it} + \rho_2 \ln P_{it-1} - \rho_3 \ln w_{it}^h + \tau_t^p + \zeta_i^p \right), \tag{16}$$

which links population—and, in turn, agglomeration—to rents, incomes, and lagged population. We include Equation (16) in an extended model ("S3")—with results shown in the fourth column of Table 3—but find little difference vis-á-vis Model D4.[33]

4.3.4 Agglomeration measure

As described in Section 3.2, we measure agglomeration, D_{it}, in terms of "effective population", where $D_{it} = P_{it} + \sum_{j \neq i} \frac{P_{jt}}{d_{ij}}$ (Holl, 2012). Notwithstanding its similarities to commonly-used market potential measures, this specification is admittedly ad-hoc.

[33] Due to the extra equation, the elpd of Model S3 is not comparable to the other models. We also find $\rho_1 = 0.089$ (s.e. 0.011), $\rho_2 = 3.9$ (s.e. 0.29), and $\rho_3 = 0.38$ (s.e. 0.10). Although we expect $\rho_3 \equiv \eta$ from Model S2, we find $\rho_3 > \eta$. This may indicate that η is negatively biased, as suggested in Footnote 32.

To understand the sensitivity of our results to the choice of agglomeration measure, we consider four alternatives for which results are presented in the last four columns of Table 3. Model S4 (column 5) uses employment, rather than population; Model S5 (column 6) uses travel-time instead of travel-distance; Model S6 (column 7) uses travel-time within an exponential decay function; and Model S7 (column 8) considers only the population of each location, ignoring those of surrounding locations.[34] Whereas Model S5 has a slightly higher elpd than Model D4, the estimated elasticities are similar in both models. For Models S6 and S7, we find lower elpd values and elasticities that are approximately one-third smaller than those for Model D4. The more limited scope of agglomeration in Models S6 and S7 is likely to capture agglomeration economies arising primarily via labour markets rather than IO linkages, which operate over longer distances. The latter, it seems, are important and better captured by the "effective population" measure.

4.3.5 Endogeneity

We test the sensitivity of our results to endogeneity. As these tests use lagged information—either as controls or instruments—we lose data from the first census in 1976. To enable like-for-like comparisons, the first column in Table 4 re-estimates Model D4 using observations from 1981–2018. In the second column ("E1"), we follow Combes, Duranton and Gobillon (2019) and include controls for two potential sources of endogeneity: Annualised population growth, G_{it}, in the prior period—to control for unobserved shocks to the labour supply—and the location quotient of highly educated workers, Q^h_{it}—to control for human capital externalities. In Models E2–E6, we address endogeneity more directly using a control function: First, we regress agglomeration on an instrument, Z_{it}, and, second, we re-estimate Model D4 including the residuals, e_{it}, from the first stage.[35]

Whereas Combes, Duranton and Gobillon (2019) use lagged population and amenity indicators as instruments, we construct a lagged Bartik shift-share instrument, Z_{it}, using data on employment for 4-digit industry sectors ($n = 64$).[36] Goldsmith-Pinkham et al.

[34] We assume travel-times to Waiheke Island equal travel-times to the centroid of Auckland, plus the travel-time by ferry and bus from Auckland to the centroid of Waiheke Island (74-minutes).

[35] Control functions have two main advantages: First, the significance of the residuals acts as a direct test for endogeneity and, second, we can allow the endogenous residuals to have non-linear effects.

[36] Let $B_{it} = \sum_k \frac{J_{kit-1}}{J_{it-1}} \frac{\hat{J}_{kt}}{\hat{J}_{kt-1}}$, where J_{kit} denotes employment in sector k, location i, at time t and $\hat{J}_{kt}$ denotes employment in sector k at time t in locations other than i. That is, B_{it} is the inner-product of lagged local sectoral shares and nationwide sectoral growth rates. We multiply $1 + B_{it}$ by lagged employment, J_{it-1}, and take logs to construct an instrument in log-levels, that is, $Z_{it} = \ln(J_{it-1}(1 + B_{it}))$.

	Model D	Endogeneity tests					
	4	E1	E2	E3	E4	E5	E6
ϵ^w	0.037	0.040	0.037	0.036	0.037	0.040	0.034
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)
ϵ^r	0.280	0.237	0.220	0.221	0.222	0.218	0.144
	(0.055)	(0.052)	(0.061)	(0.061)	(0.059)	(0.053)	(0.054)
E_p	0.051	0.049	0.045	0.045	0.045	0.047	0.036
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)
E_c	0.031	0.018	0.017	0.018	0.018	0.013	0.001
	(0.014)	(0.014)	(0.016)	(0.016)	(0.015)	(0.014)	(0.014)
Controls	No	Yes	No	No	No	Yes	No
C.F.	No	No	Yes	Yes	Yes	Yes	Yes
Rsq. [%]	99.1	99.1	99.1	99.1	99.1	99.1	99.1
elpd	1,989	1,985	2,002	1,998	1,987	1,989	2,028

Table 4: Estimates for models described in Section 4.3.4 (s.e. in parentheses). All models use 1,059 observations and include individual year and location effects. Models E1 and E5 control for (1) annualised population growth in the preceding period and (2) the location quotient for highly educated workers. Models E2–E6 first regress agglomeration on a Bartik shift-share instrument and, second, include the residuals, e_{it}, from the first step in the model. The residuals e_{it} enter linearly in Models E2 and E5, as the argument to a GAM in Model E3, and with measurement error in Model E4. Model E6 uses commute-weighted Bartik instrument, $Z^s_{it} = \sum_j x_{jt|i} Z_{jt}$.

(2020) argue the validity of Z_{it} hinges on the exogeneity of lagged local sectoral shares, although Borusyak et al. (2022) demonstrate Z_{it} is valid when shocks are as-good-as-randomly assigned—conditional on their unobservable elements. In our setting, we suggest national employment growth rates act as randomly-assigned labour demand-shifters across industry sectors, such that Z_{it} is valid.[37] Figure 5 plots Z_{it} (horizontal axis) versus agglomeration (left panel), $\ln D_{it}$, and the residuals from the wage (middle panel) and rent (right panel) equations in Model D4, respectively. Promisingly, we observe a strong association between $\ln D_{it}$ and Z_{it} but no association between the latter and $\ln w_{it}$ or $\ln r_{it}$—conditional on other explanatory variables in Model D4.

As described in the notes to Table 4, Models E2–E5 differ in the way the residuals, e_{it}, from the first stage enter the second stage. In contrast, Model E6 replicates Model E2 using a commute-share weighted instrument, $Z^s_{it} = \sum_j x_{jt|i} Z_{jt}$. We find evidence of endogeneity, with precisely estimated coefficients for e_{it} in Models E2–E6. Compared to Model D4, we find similar estimates of the income elasticity, ϵ^w, but smaller estimates of the rent elasticity, ϵ^r.[38] In Model E6, the magnitude of ϵ^r, for example, is approximately half that

[37] The local sectoral shares may also be exogenous, as census data introduces lags of at least five years and the spatial income premia control for sectoral composition at the two-digit level, among other characteristics.
[38] Ex ante, we expect endogeneity is a greater risk to ϵ^r, as the controls used to derive the income premia are more comprehensive than those for the rent premia. See Section 3.1 for details.

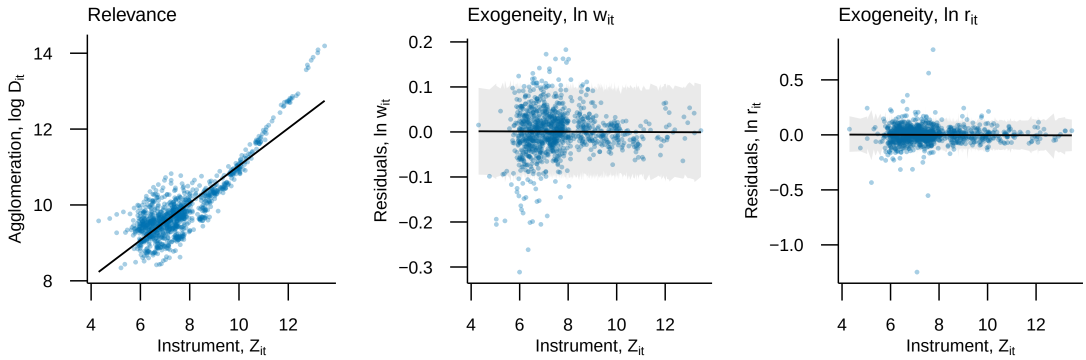

Figure 5: Left, middle, and right panel plot the instrument (horizontal axis), Z_{it}, versus agglomeration, $\ln D_{it}$, and median residual from the income and rent equation, respectively. In the middle and right panel, we show the median and associated 95% credibility interval estimate from a regression of the residuals versus Z_{it}, where we allow for uncertainty in the estimates of the residuals.

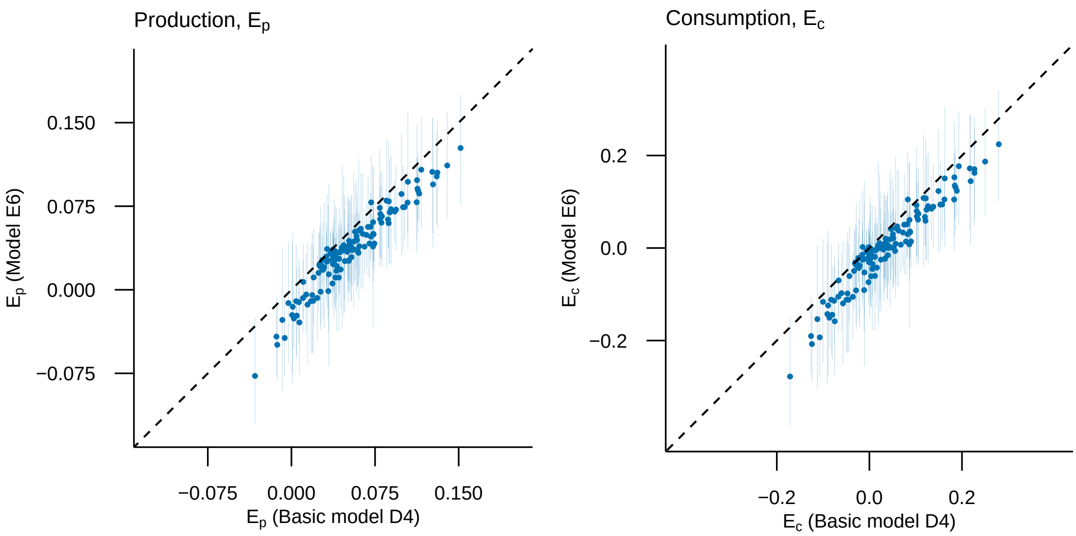

Figure 6: Location-specific estimates of E_p (left panel) and E_c (right panel) from Model D4 (horizontal axis) and Model E6 (vertical axis). The dashed 45 degree line denotes the points where both models yield the same estimates and the error bars denote 95% credible intervals for Model E6.

of Model D4, although there remains considerable overlap in the distribution of estimates. This is reinforced by Figure 6, which reveals small differences in the location-specific estimates of agglomeration economies between Models D4 and E6. Finally, we note our estimates for ϵ^r are close to those in Combes, Duranton and Gobillon (2019), which uses panel data to estimate the elasticity of urban house prices with respect to population in France and report 0.176–0.304 for pooled OLS (c.f. Table 4) and 0.215–0.267 for instrumental variables (c.f. Table A1). The alignment in elasticities of housing costs is reassuring given large differences in contexts, methods, and instruments.

5 Discussion

The results in Section 4 suggest Model D4 produces estimates of agglomeration economies that are quantitatively similar to other models, such as the Commuting model (c.f. Section 4.2) and models embodying various sensitivity tests (c.f. Section 4.3). For this reason, we use Model D4 as the basis for our discussion while noting that the key findings carry over to other models. To begin, we compare estimates derived from aggregate data (Model A) with micro-data (Model D4).[39] Interestingly, the differences in estimates derived from aggregate and micro-data are broadly constant in the period 1976–2018. Like Combes, Duranton and Gobillon (2008), we have evidence that "sorting matters", even if its effects appear to be relatively stable over time—at least in the New Zealand context.

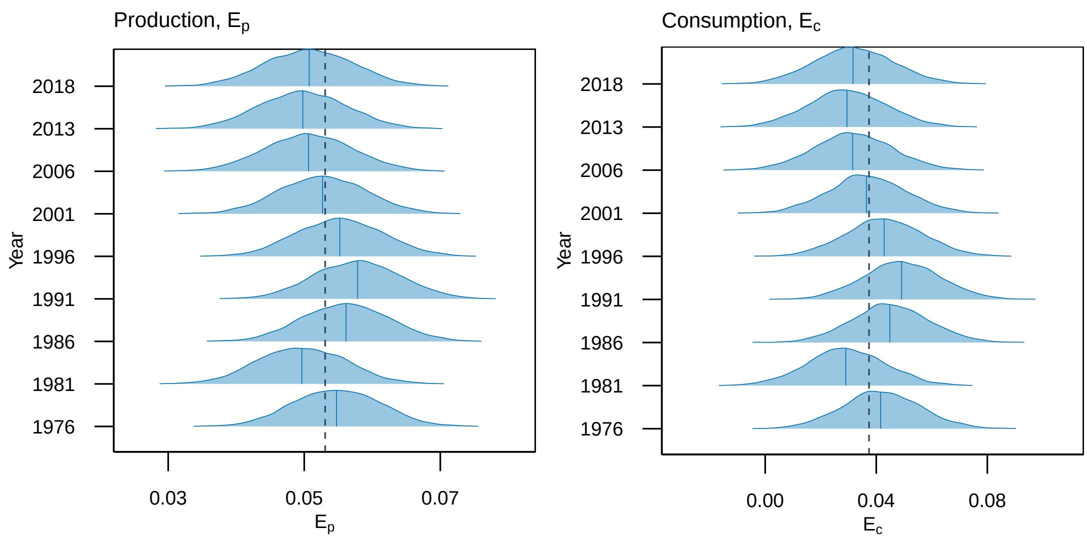

Figure 7: Trends in agglomeration economies in production (left panel) and consumption (right panel), as derived from Model D4. The dashed vertical lines denote the median for the sample.

Figure 7 shows trends in estimates of agglomeration economies in production, E_p (left panel), and consumption, E_c (right panel), for Model D4. We find subtle temporal variation: From a nadir in 1981, estimates peak in 1991, and then gradually decline. Between 1991 and 2006, E_p and E_c fell by approximately 0.7 and 1.4 percentage points, respectively, but have since been relatively stable. Although the magnitude of this decline is similar to that found in the meta-analysis by Donovan et al. (2021), the peak in the latter study occurs approximately one decade after that which we find here. Informally,

[39] We estimate an equivalent version of Model D4 using aggregate data, although without errors-in-outcomes. Results for this model are not reported but are available on request from the authors.

we observe less temporal variation for the three most recent Censuses compared to those in the period 1976–2001. To explain trends, Donovan et al. (2021) points to general factors, such as ICT, which may have affected the economic advantages of cities in this period. The trends we observe here may also have a domestic explanation: In the 1980s New Zealand implemented major economic policy reforms (Evans et al., 1996). These reforms included, but were not limited to, the complete removal of agricultural subsidies, which reduced the financial returns from agriculture and led to a collapse in the price of agricultural land (Vitalis, 2007). Greater temporal variation in the period 1976–2001 may reflect the lingering effects of these policy shocks.[40] Notwithstanding this historical variation, our recent estimates of agglomeration economies are relatively stable.

In contrast, we find more significant spatial variation in agglomeration economies. Figure 8 maps median estimates of agglomeration economies in production, E_p , and consumption, E_c, for each location. We observe some clustering in the results: Locations that are close to larger cities—such as Auckland (population 1.4 million), Wellington (population 404,000), and Christchurch (population 385,000)—tend to have stronger agglomeration economies, and vice versa for more remote locations. To provide more detailed insight into this variation, Figure 9 plots E_p (vertical axes) versus E_c (horizontal axes) for the six cities with the largest populations. In these panels, each point represents an individual posterior estimate, whereas the solid lines indicate the median for each city and the dashed lines indicate the median for the sample. For these cities, we find estimates of E_p and E_c of approximately 0.04 and 0.00, respectively, which is smaller than but close to the median for locations in the sample. In New Zealand's largest cities, agglomeration appears to benefit production but not consumption.[41] Broad similarities for larger cities belies more notable differences for smaller locations. In Figure 10 and Figure 11, we plot the six locations with the largest and smallest agglomeration economies, respectively, as measured by the geometric average of their medians. We find estimates for E_p and E_c vary from approximately 0.15 and 0.20 in Figure 10 to 0.00 and -0.10 in Figure 11, respectively. All the locations in Figure 10 and Figure 11 are relatively small, with populations less than 7,000. On the other hand, the locations in Figure 10 are all proximate to large cities whereas those in Figure 11 are relatively remote.[42]

[40] This explanation has parallels to recent legal scholarship by Sitaraman et al. (2021), which argues deregulatory initiatives in the U.S. during the 1980s and 1990s—for example, in transport, communications, trade policy, and anti-trust and corporate consolidation—exacerbated urban and rural inequalities. Major economic policy reforms can, according to this thesis, have spatial dimensions.

[41] Morrison (2011) finds reported well-being is lower than average in New Zealand's larger urban areas.

[42] Figure 11 also reveals hints of clustering, with four of the six locations—namely, Kaitaia, Paihia, Kawakawa, and Kaikohe—located in the Far North District and the other two located in the central North Island.

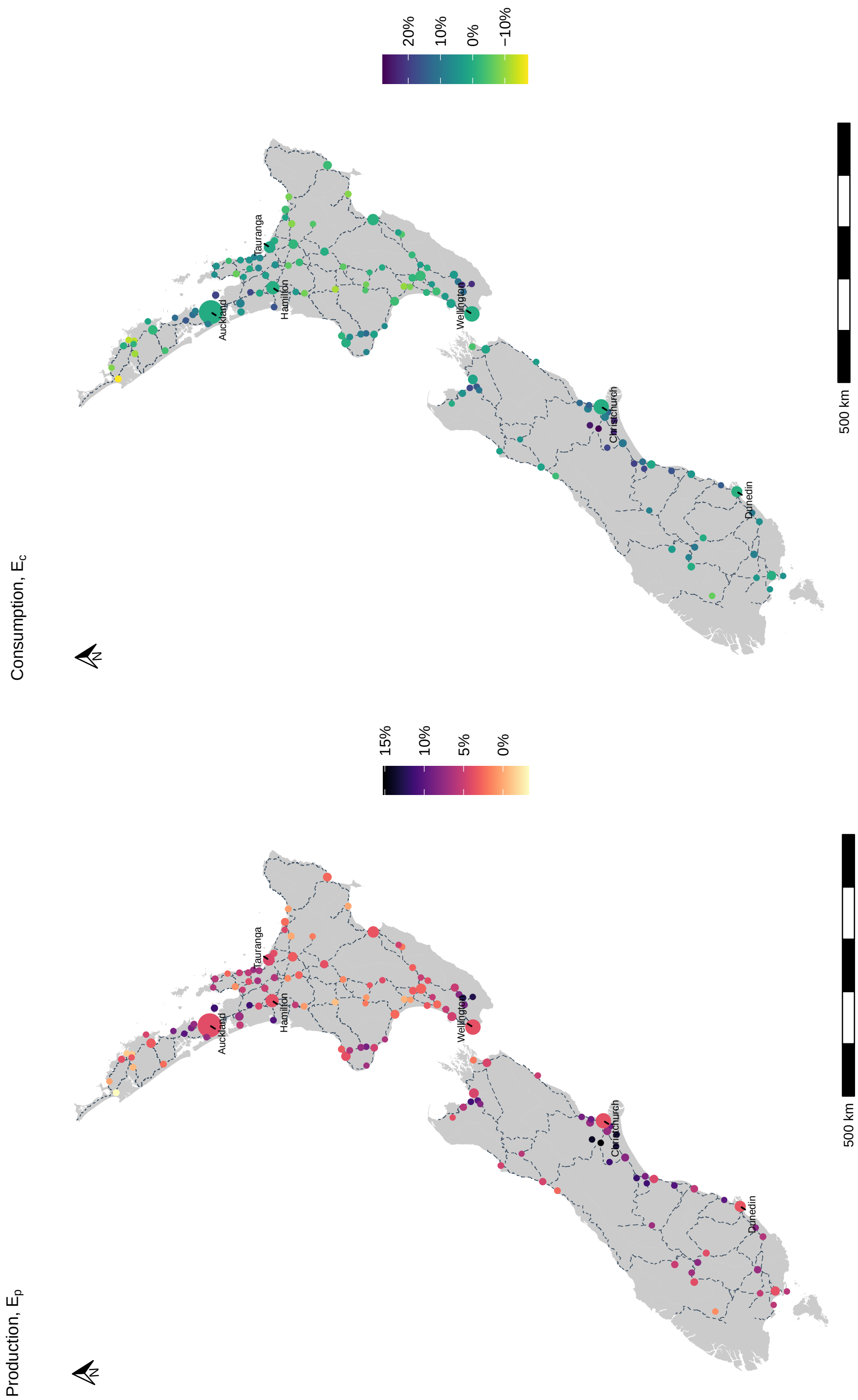

Figure 8: Agglomeration economies in production (bottom) and consumption (top). These are the median estimate for each location, as per Model D4. Labels denote the six locations with the largest populations, whereas dashed lines denote highways.

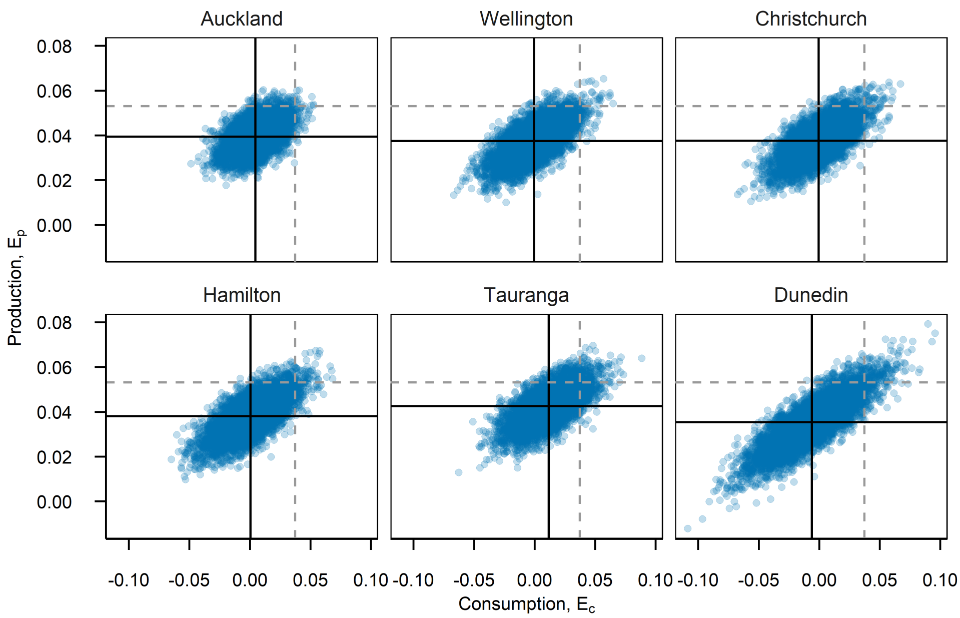

Figure 9: Estimates of E_p (vertical axes) and E_c (horizontal axes) for the largest cities. Dashed and solid lines indicate median for the sample and location, respectively.

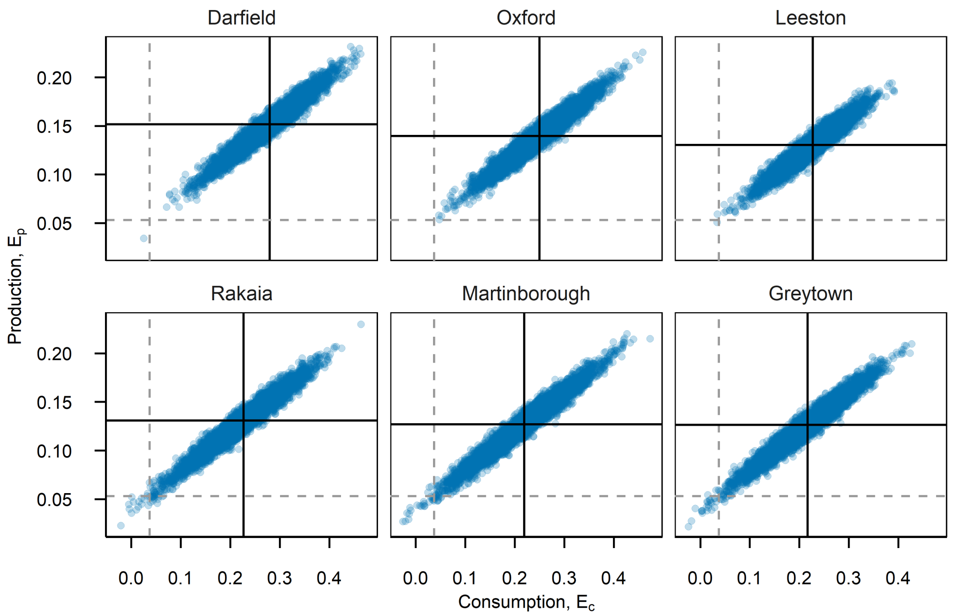

Figure 10: Estimates of E_p (vertical axes) and E_c (horizontal axes) for locations with the largest average effects. Dashed and solid lines indicate median for the sample and location, respectively.

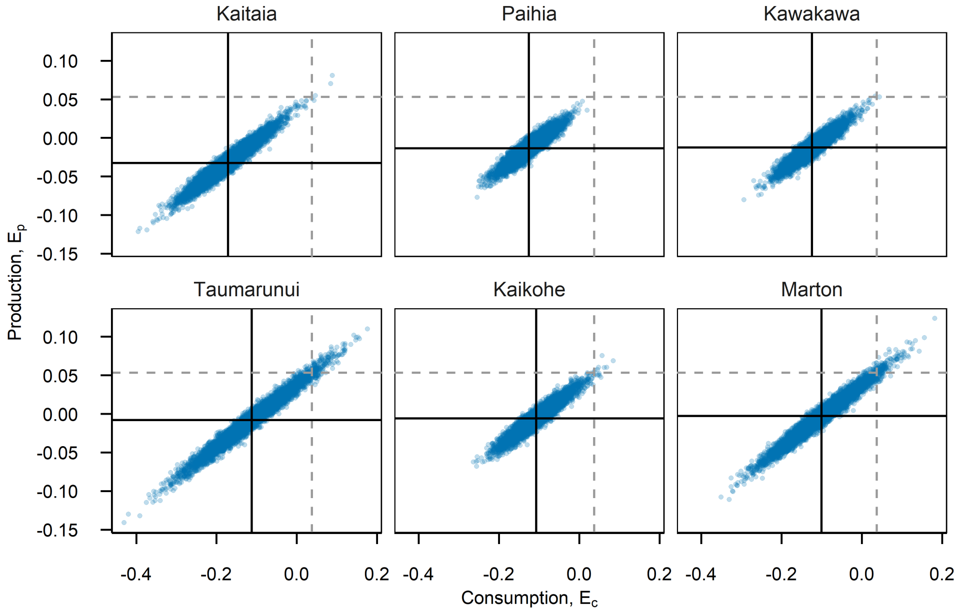

Figure 11: Estimates of E_p (vertical axes) and E_c (horizontal axes) for locations with the smallest average effects. Dashed and solid lines indicate median for the sample and location, respectively.

To proceed, we define three types of locations based on 2018 populations: "Satellite A" describes locations that are within 90-minutes travel-time of cities with populations of at least 300,000; "Satellite B" describes locations that are within 45-minutes travel-time of cities with populations of at least 50,000; and "Remote" describes locations that are more than 60-minutes travel-time of cities with populations of at least 50,000.[43] Figure 12 relates these definitions to estimates of agglomeration economies in production, E_p. The top left panel plots E_p for each location (vertical axis) versus agglomeration (horizontal axis), $\ln D_{it}$, where the error bars indicates the 90% credible interval. Heterogeneity between locations exhibits a familiar "funnel" pattern, with E_p converging towards the median as $\ln D_{it}$ increases. Uncertainty within locations also reduces with $\ln D_{it}$, possibly due to the precision of the spatial premia. The top-right, bottom-left, and bottom-right panels in Figure 12 then highlight the E_p for "Satellite A", "Satellite B", and "Remote", respectively. On average, the E_p for "Satellite A" and "Satellite B" locations lie above the median, and vice versa for "Remote" locations. To see whether these definitions have any bite, we extend Model D4 to include a group-level intercept per location-type as well as an interaction between location-type and $\ln D_{it}$. We find that model performance improves

[43] In 2018, only three cities—namely, Auckland, Wellington, and Christchurch—had populations greater than 300,000 (Satellite A), whereas 17 cities had populations greater than 50,000 (Satellite B).

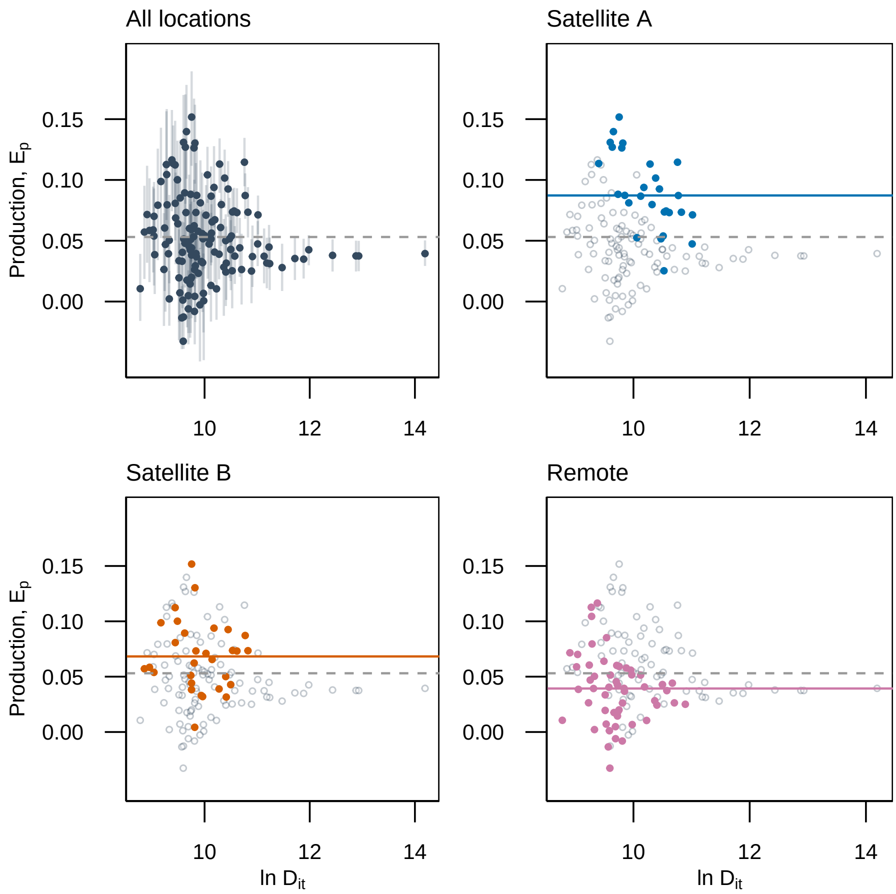

Figure 12: Top-left panel: Estimates of E_p for each location (vertical axis) versus agglomeration (horizontal axis), $\ln D_i$. The top-right, bottom-left, and bottom-right panels show the median estimates for "Satellite A", "Satellite B", and "Remote" locations, respectively, versus the median for the sample. Locations that meet the definition for both Satellite A and B are shown in both panels.

and "Satellite A" towns have larger rent elasticities. We posit these locations may offer some of the benefits of large cities, such as access to airports, with less congestion.

To finish, we note three limitations of our analysis. First, we lack some data, such as commuting shares, $x_{jt|i}$, historical transport costs, and housing attributes—such as floor space and land—which may bias our estimates of commute costs, $\ln \hat{C}_{it}$, and the spatial rent premia, $\ln r_{it}$. Second, we find that rent is inelastic with respect to income, which may indicate the need for a more flexible model of worker preferences.[44] Third, we do not allow for spatial variation in the price of the composite consumption good, Y_{it}, which may impart a downwards bias to estimates of agglomeration economies in large cities.

[44] Including income in the rent equation (c.f. Section 4.3.2), however, does not change our results.

6 Conclusion

When thinking about agglomeration economies, our results suggest the forking paths analogy needs an extra dimension: A multiplicity of outcomes can arise over both time *and* space. That said, the temporal variation we find is relatively subtle: From a nadir in 1981, estimates peak in 1991, and then gradually decline—with agglomeration economies in production and consumption falling by approximately 0.7 and 1.4 percentage points, respectively, between 1991 and 2006. Since this time, however, our estimates of agglomeration economies remain relatively stable; contrary to popular claims, technology does not appear to have made the world "flatter". In contrast, we find evidence of more notable spatial variation in agglomeration economies: Larger cities in New Zealand have advantages in production but not in consumption, whereas small towns that are close to large cities, or "satellites", experience agglomeration economies that are stronger than average. Taken together, our results suggest that agglomeration economies are rather stable—at least in recent decades—but cast some doubt on their transferability—at least domestically within New Zealand. The latter is especially relevant to smaller towns, where our results reveal a wide range of possible outcomes, or "forking paths."

To the extent that our findings raise questions over the stability and transferability of agglomeration economies, they also highlight the need for further research and nuanced policies. More research is needed to trace the evolution of agglomeration economies within countries and over time and, where possible, explain the resulting variation in estimates. We also repeat calls for more research into urban advantages in both production and consumption. Although agglomeration seems to enhance the productivity of cities and towns in New Zealand, we find more tenuous benefits for consumption. Whether this reflects the inevitable erosion of amenity due to congestion effects, or a failure to adopt effective policies to mitigate congestion, remains an open question. In terms of empirical measurement, we see merit in exploring agglomeration indicators that respond to multi-modal transport costs—rather than simply car travel-time or distance. In New Zealand, for example, the combination of international remoteness, complex geography, and lack of inter-city passenger rail services may increase the relative importance of access to airports. Finally, we underscore the need for nuanced urban policies. For example, although we find evidence of stronger agglomeration economies in satellite locations, the latter typically have small populations and, presumably, less congestion—both of which are endogenous attributes. Policies would, ideally, recognise and respond to these temporal and spatial forks in the road.

References

Ahlfeldt, G. M. and E. Pietrostefani (2019). 'The economic effects of density: A synthesis'. *Journal of Urban Economics* 111, pp. 93–107.

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm and N. Wolf (2015). 'The Economics of Density: Evidence from the Berlin Wall'. *Econometrica* 83.6, pp. 2127–2189.

Ahrend, R., E. Farchy, I. Kaplanis and A. C. Lembcke (2017). 'What makes cities more productive? Evidence from five OECD countries on the role of urban governance'. *Journal of Regional Science* 57.3, pp. 385–410.

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

Autant-Bernard, C. and J. P. LeSage (2011). 'Quantifying knowledge spillovers using spatial econometric models'. *Journal of Regional Science* 51.3, pp. 471–496.

Borges, J. L. et al. (1962). 'The garden of forking paths'. *Labyrinths: selected stories and other writings*, pp. 19–29.

Borusyak, K., P. Hull and X. Jaravel (2022). 'Quasi-experimental shift-share research designs'. *The Review of Economic Studies* 89.1, pp. 181–213.

Bürkner, P.-C. (2017). 'brms: An R package for Bayesian multilevel models using Stan'. *Journal of Statistical Software* 80.1, pp. 1–28.

Cairncross, F. (1997). *The death of distance: How the communications revolution will change our lives*. Vol. 302. Harvard Business School Press Boston, MA.

Chauvin, J. P., E. L. Glaeser, Y. Ma and K. Tobio (2017). 'What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States'. *Journal of Urban Economics* 98, pp. 17–49.

Chen, X., C. Xuan and R. Qiu (2021). 'Understanding spatial spillover effects of airports on economic development: New evidence from China's hub airports'. *Transportation Research Part A: Policy and Practice* 143, pp. 48–60.

Combes, P.-P., G. Duranton and L. Gobillon (2008). 'Spatial wage disparities: Sorting matters!' *Journal of Urban Economics* 63.2, pp. 723–742.

— (2019). 'The costs of agglomeration: House and land prices in French cities'. *The Review of Economic Studies* 86.4, pp. 1556–1589.

Combes, P.-P., G. Duranton, L. Gobillon and S. Roux (2010). 'Estimating Agglomeration Economies with History, Geology, and Worker Effects'. *Agglomeration Economics*. University of Chicago Press, pp. 15–66.

Donovan, S., T. de Graaff, H. L. F. de Groot and C. Koopmans (2021). *Unravelling urban advantages—A meta-analysis of agglomeration economies*. Working Paper No. 21-026/VIII. Amsterdam, The Netherlands: Tinbergen Institute.

Ehrl, P. (2013). 'Agglomeration economies with consistent productivity estimates'. *Regional Science and Urban Economics* 43.5, pp. 751–763.

Evans, L., A. Grimes, B. Wilkinson and D. Teece (1996). 'Economic reform in New Zealand 1984-95: The pursuit of efficiency'. *Journal of Economic Literature* 34.4, pp. 1856–1902.

Fabling, R. and D. C. Maré (2019). *Improved productivity measurement in New Zealand's Longitudinal Business Database*. Tech. rep. Motu Economic and Public Policy Research.

Fosgerau, M. and M. Bierlaire (2009). 'Discrete choice models with multiplicative error terms'. *Transportation Research Part B: Methodological* 43.5, pp. 494–505.

Friedman, T. L. (2006). *The world is flat: The globalized world in the twenty-first century*. Penguin London.

Gabriel, S. A. and S. S. Rosenthal (2004). 'Quality of the business environment versus quality of life: Do firms and households like the same cities?' *Review of Economics and Statistics* 86.1, pp. 438–444.

Gaigné, C., H. R. Koster, F. Moizeau and J.-F. Thisse (2022). 'Who lives where in the city? Amenities, commuting and income sorting'. *Journal of Urban Economics* 128, p. 103394.

Gautier, P. A., M. Svarer and C. N. Teulings (2010). 'Marriage and the city: Search frictions and sorting of singles'. *Journal of Urban Economics* 67.2, pp. 206–218.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari and D. B. Rubin (2013). *Bayesian data analysis*. CRC press.

Gelman, A. and J. Hill (2007). *Data analysis using regression and multilevel hierarchical models*. Vol. 1. Cambridge University Press New York, NY, USA.

Gelman, A. and E. Loken (2013). 'The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time'. *Department of Statistics, Columbia University* 348.

Geweke, J. (1993). 'Bayesian treatment of the independent Student-t linear model'. *Journal of Applied Econometrics* 8.S1, S19–S40.

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

— (2011). *Triumph of the City*. Pan.

Glaeser, E. L., J. Gyourko, E. Morales and C. G. Nathanson (2014). 'Housing dynamics: An urban approach'. *Journal of Urban Economics* 81, pp. 45–56.

Goldsmith-Pinkham, P., I. Sorkin and H. Swift (2020). 'Bartik instruments: What, when, why, and how'. *American Economic Review* 110.8, pp. 2586–2624.

Groot, S. P. T., H. L. F. de Groot and M. J. Smit (2014). 'Regional wage differences in the Netherlands: Micro evidence on agglomeration externalities'. *Journal of Regional Science* 54.3, pp. 503–523.

Holl, A. (2012). 'Market potential and firm-level productivity in Spain'. *Journal of Economic Geography* 12.6, pp. 1191–1215.

Kasy, M. (2021). 'Of Forking Paths and Tied Hands: Selective Publication of Findings, and What Economists Should Do about It'. *Journal of Economic Perspectives* 35.3, pp. 175–92.

Maré, D. C. and D. J. Graham (2013). 'Agglomeration elasticities and firm heterogeneity'. *Journal of Urban Economics* 75, pp. 44–56.

Maré, D. C. and J. Poot (2019). *Valuing cultural diversity of cities*. Working Paper. Motu Economic and Public Policy Research.

Martínez-Galarraga, J., E. Paluzie, J. Pons and D. A. Tirado-Fabregat (2008). 'Agglomeration and labour productivity in Spain over the long term'. *Cliometrica* 2.3, pp. 195–212.

MBIE (2015). *Regional Economic Activity Report 2015*. Tech. rep. Ministry of Business, Innovation and Employment.

Meier, S. and C. D. Sprenger (2015). 'Temporal stability of time preferences'. *Review of Economics and Statistics* 97.2, pp. 273–286.

Morrison, P. S. (2011). 'Local expressions of subjective well-being: The New Zealand experience'. *Regional Studies* 45.8, pp. 1039–1058.

Puga, D. (2010). 'The magnitude and causes of agglomeration economies'. *Journal of Regional Science* 50.1, pp. 203–219.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rice, P., A. J. Venables and E. Patacchini (2006). 'Spatial determinants of productivity: Analysis for the regions of Great Britain'. *Regional Science and Urban Economics* 36.6, pp. 727–752.

Roback, J. (1982). 'Wages, rents, and the quality of life'. *Journal of Political Economy* 90.6, pp. 1257–1278.

Rosenzweig, M. R. and C. Udry (2020). 'External validity in a stochastic world: Evidence from low-income countries'. *The Review of Economic Studies* 87.1, pp. 343–381.

RStudio Team (2021). *RStudio: Integrated Development Environment for R*. Boston, MA.

Shenker, J. (2020). 'Cities after coronavirus: how Covid-19 could radically alter urban life'. *The Guardian*.

Sitaraman, G., M. Ricks and C. Serkin (2021). 'Regulation and the Geography of Inequality'. *Duke Law Journal* 70.8, pp. 1763–1836.

Stats NZ (2017). *Statistical standard for geographic areas 2018*. Tech. rep. Statistics New Zealand.

Vehtari, A., A. Gelman and J. Gabry (2017). 'Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC'. *Statistics and Computing* 27.5, pp. 1413–1432.

Vitalis, V. (2007). 'Agricultural subsidy reform and its implications for sustainable development: the New Zealand experience'. *Environmental Sciences* 4.1, pp. 21–40.

Waka Kotahi (2020). *Monetised Benefits and Costs Manual*. Technical Report Version 1.5. The New Zealand Transport Agency.

Zellner, A. and T. Ando (2010). 'A direct Monte Carlo approach for Bayesian analysis of the seemingly unrelated regression model'. *Journal of Econometrics* 159.1, pp. 33–45.