Infrastructure's Long-Lived Impact on Urban Development: Theory and Empirics

Arthur Grimes,[1,2] Eyal Apatov,[2] Larissa Lutchman,[2] Anna Robinson[1]

[1] Motu Economic and Public Policy Research, Wellington, New Zealand
[2] University of Auckland, Auckland, New Zealand

Presenting author: Arthur Grimes
Motu Economic & Public Policy Research
PO Box 24390, Wellington 6142
arthur.grimes@motu.org.nz

Keywords: Infrastructure, city development, population growth, migration, spatial equilibrium

JEL Nos.: H54, R12

Abstract

We analyse impacts that infrastructure provision and other factors have on long run urban growth. Reflecting a spatial equilibrium approach, growing cities have preferred attributes relative to other cities. Social and transport infrastructure have both productive and amenity value and so may enhance a city's growth. We outline a new theoretical model that includes distance-related effects on individual utility and thence population location, and we test this model using historical data covering 1926 to 2006 across 56 New Zealand towns. Instruments dating back to 1880 are used to deal with potential endogeneity issues, and we use spatial-econometrics techniques to test for spatial spillovers between cities.

Acknowledgements

This draft paper has been prepared with funding from the Resilient Urban Futures programme sourced from the Ministry of Business Innovation and Employment (New Zealand). We thank all public agencies (especially Statistics New Zealand) and private sources from whom we accessed historical data.

1. Introduction

What effects do infrastructure investments and other factors have on long term urban development? We address this question using a newly specified theoretical model and using data covering 80 years across 56 New Zealand towns. The long run city population data and data for many of our other variables have been derived from Statistics New Zealand Yearbooks dating back to the start of the twentieth century.[1] The determinants of town and city development is of importance to policy-makers when deciding whether, and where, to invest in major transport and social infrastructure projects. The analysis helps policy-makers to understand the intended, and potentially unintended, long run consequences of their infrastructure investment decisions.

Reflecting a spatial equilibrium approach (Overman et al, 2010; Grimes, 2014), we maintain that population flows reflect people's overall rankings of urban areas. Thus, through revealed preference, growing cities are shown to have preferred attributes (wages and amenities combined, adjusted for costs) relative to other cities. Social infrastructure (such as higher educational institutions and hospitals) and transport infrastructure may have both productive and amenity value. Thus increased provision of such infrastructure within a city may enhance a city's attractiveness provided that the benefits of the new infrastructure exceed local costs of provision. Agglomeration benefits may magnify the benefits of infrastructure investments, especially in larger cities. Poor infrastructure provision linking an urban area to major cities and other amenities may, conversely, reduce the attractiveness of that urban area, curtailing its long run population growth. In the next section, we summarise key insights gained from prior studies about the effects of infrastructure investments on city development. Two specific areas are highlighted – the effects of transport infrastructure and the effects of higher educational institutions – to illustrate effects of infrastructure assets that have differing mixes of productive and amenity value.

We then outline a theoretical model that includes distance-related effects on individual utility, incomes and costs. Ceteris paribus, people favour living close to amenities, and they earn higher wages when they are located in or near a major agglomeration. Enjoyment of amenities declines as distance to those amenities increases, and wages decline as distance from the major agglomeration increases. Transport costs increase as distance to these assets increases. Each of these factors influences urban population growth. The model is related to that in a recent paper by Duranton and Turner (2012). However, the new specification avoids a convenient but questionable assumption in their approach in relation to the effect of distance on individual utility.

We test our model using a newly derived long-term (80 year) historical series on urban populations measured every 10 years from 1926 to 2006 for 56 towns across New Zealand. This dataset enables us to relate population growth of these urban areas to early and subsequent infrastructure provision and to initial conditions (e.g. existence of a harbour, topography, climate, etc). We include tests of the impacts of a number of social and transport infrastructure variables. In addition, unobserved social amenities and wages are hypothesised to be related positively to existing population levels. Non-infrastructure control variables include climatic variables, land-use capability, regional variables and a human capital measure.

[1] The digitisation of all Statistics New Zealand Yearbooks has been of invaluable assistance in the derivation of long term data for this analysis. All digitised Yearbooks are available at: http://www.stats.govt.nz/browse_for_stats/snapshots-of-nz/digital-yearbook-collection.aspx.

Some of our measures date back to the 1880s in order to account for endogeneity and to model the long-lived effects of certain conditions. For instance, we have data for the ratio of Maori to total population in 1881 and we use data obtained from the 1880 Railway Commission report. Additional variables obtained from early Statistics New Zealand Yearbooks serve as instruments for potentially endogenous variables in our regressions. As well as dealing with endogeneity issues through our choice of instruments, we test for the presence of random effects and for spatial spillovers between cities.

The empirical analysis shows that five dominant factors have impacted positively on urban growth, especially since 1966: land-use capability, human capital, sunshine hours, population size and proximity to the country's dominant city, Auckland. In our concluding section, we interpret how these results may usefully influence the formulation and implementation of infrastructure policy.

2. Prior Literature

2.1 Infrastructure and Population Growth

Models of spatial equilibrium demonstrate how population flows across regions in order to equate utility in different areas (Glaeser and Gottlieb, 2009; Overman et al, 2010; McCann, 2013). In these models, individual utility is derived from consumption of amenities plus private consumption of tradable and non-tradable goods (where the price of the former is exogenous to the region and that of the latter is endogenous). Consumption is restricted by the individual's budget constraint where wages may be city-specific reflecting agglomeration and other factors.

Grimes (2014) extends the Overman et al model to include infrastructure provision deriving the conditions under which a new infrastructure investment within a city will expand that city's population. To do so, the infrastructure investment must raise amenity-adjusted real wages, where amenity-adjusted wages include the value of unpriced amenities to an individual. An infrastructure investment may increase amenity-adjusted wages through a variety of mechanisms: first, the infrastructure may raise amenities in a city (e.g. through provision of a new concert hall); second, the infrastructure may reduce travel costs (e.g. through provision of an improved transport network); third, the infrastructure may raise productivity and hence wages (e.g. through a new port or airport); fourth, the infrastructure may raise skills and hence wages (e.g. through provision of a higher educational institution). However, the new infrastructure may result in cost increases, for instance through higher taxes to pay for the new facilities and through higher land costs (house prices) as new population is attracted to the city. The latter effect, which occurs as a result of net inward migration in response to the new investment, is the mechanism by which the spatial adjustment to the new infrastructure is equilibrated.

Empirical applications of the spatial equilibrium approach can be separated into those that deal with localised infrastructure (within a locality) and those that deal with infrastructure connecting cities. An example of the former is the study by Duflo and Pande (2007) of the localised impact of the construction of dams in India. An example of the latter is the study by Coleman (2012) of the effect of the construction of the Erie Canal on economic activity in rural areas of New York state. Another example is that of Gibbons et al (2012) who examine the effects of new inter-city road infrastructure on firm outcomes in the UK. Each of these studies uses an exogenous event (construction of a dam, canal or inter-city road) to examine economic outcomes. Where such an event is not available, careful testing has to be undertaken to ensure that the infrastructure that is the subject of study is not an endogenous response to population growth. Where it may be an endogenous response, the use of exogenous instruments in estimation (as in Wu and Gopinath, 2008) is required.

2.2 Transport Infrastructure and Regional Growth[2]

Early studies which find positive impacts of transport infrastructure on economic growth include Mera (1973) for Japan's regions, Blum (1982) for regional growth in West Germany, and Aschauer (1989) and Munnell (1990) for regions within the United States. Economic growth induced by transport investments encourages employment and population growth as consumers move across regions to maximise wages. Thus transport investments result in population growth and employment growth within regions where imperfect, spatially competitive labour markets lead to the provision of higher net wages (Fujita and Thisse, 2002). Early spatial infrastructure studies, however, tended to suffer from a lack of attention to the potential endogeneity of transport links.

Population changes within metropolitan areas and employment growth across metropolitan areas have been the focus of more recent analysis of the role that the United States interstate highway system has played in the development of cities (Baum-Snow, 2010; Duranton and Turner, 2012). Both studies estimate the effect of state highway infrastructure on regional population growth and share the same main instrumental variable, the 1947 plan of the US interstate highway system to account for the potential endogeneity of the highway network.[3] However, the foci of the investigations differ. Duranton and Turner explore the long term effect of transport infrastructure on regional population growth, whereas Baum-Snow examines its impact on within city population decentralisation. Baum-Snow finds that highways lead to people residing within suburban areas rather than within the central city, and that declining city transport costs as a result of road construction has led to firm productivity gains, resulting in higher wages for workers. Duranton and Turner's analysis finds that a 10% increase in a given city's stock of interstate highways leads to a 1.5% increase in employment over 20 years.

Using similar instruments, Duranton et al (2013) find that the quality of the highway network affects the structure of a city's production, with a 10% increase in a city's highways leading to a 5% increase in tonnes of goods exported by that city. This result mirrors earlier results on the importance of the transportation network for city production structures (Fernald, 1999). Similarly, the quality of the transportation network may affect the degree of agglomeration economies within and surrounding a city (Fujita and Thisse, 2002; McCann, 2013; Maré and Graham, 2013). However,

[2] Lutchman (2013) provides an in-depth discussion of the relevant literature.
[3] Duranton and Turner argue that instrumental variables for road and highway networks must also control for historical population since historical population levels affect future population growth independently from highway infrastructure.

improved transportation links do not necessarily lead to agglomeration for all sectors. Glaeser (1998) suggests that declining transport costs within the United States led to fewer jobs within the manufacturing sector within cities that have high urban densities, while Behrens and Picard (2011) find that freight rate differentials can incentivise manufacturing firms to scatter across space instead of clustering. Service sectors benefit from falling transport costs through the benefits of clustering, and thus choose to locate within cities. In their study of the distance decay of agglomeration benefits, Graham et al (2009) conclude that both the distance decay and productivity impacts of agglomeration are relatively greater for firms in services than for those in manufacturing.

Beyond its contribution to production, transport infrastructure has value by reducing costs for consumers who reside within close proximity to it. If consumers or firms prefer to locate within close proximity to these interchanges, their demand will be reflected in the increased price of housing or commercial buildings in the immediate area (Haughwout, 2002). Transport corridors that are able to deliver both mobility and amenity improvements have been found to deliver improved economic outcomes reflected in increased land rents (Donovan and Munro, 2013; Grimes and Liang, 2010, Grimes and Young, 2013).

One issue in modelling the impacts of transport infrastructure is the potential need to take into account spatial spillovers. Evidence for the existence of regional spillovers related to transport infrastructure is mixed and may depend on the definition and size of 'regions'. Neither Holtz-Eakin and Schwartz (1995) nor Duranton and Turner (2012) finds statistically significant spillover effects of highways across regions in the United States. By contrast, a general method of moments (GMM) estimate of a dynamic regional production function that includes the spillover effects of highways in US states finds that neighbouring states acquire some of the productivity benefits of highway improvements carried out in a nearby state (Jiwattanakulpaisarn et al, 2011). Similarly, within China, Yu et al (2013) find that land transport investment in neighbouring regions has a significant spillover effect across regions but the magnitude of the effect differs depending on the current productivity of the regional economy. Ding (2013) supports these propositions with analysis of the positive spillover effects associated with urban roads and regional roads for Chinese regions.

Of the above studies, Duranton and Turner's investigation of transport infrastructure and regional growth is the most similar to ours. Their model specification originates directly from consumer theory, with the inclusion of variables for distance travelled and exogenous amenities within a city in the representative resident's utility function. This approach yields equations for three variables: the rate of change of population, investment in roads, and initial road characteristics. Population change is a function of the prior period's level of population and roading, plus observable time-invariant regional characteristics. Investment in roads is a function of the same variables while initial road characteristics are a function of the prior population level, observable time-invariant regional characteristics and a vector of exogenous (historical) regional characteristics. However, Duranton and Turner's postulated consumer utility function treats distance travelled by an individual as contributing positively to consumer utility which contrasts with the notion that travel is a cost. Our theoretical approach uses that of Duranton and Turner as a starting point but instead treats distance travelled as a negative contribution to utility in keeping with the more standard treatment of distance as a cost.

2.3 Higher Educational Institutions, Skills and Regional Growth[4]

The impact of Higher Education Institutions (HEIs) on regional growth can be interpreted within the context of endogenous growth models which relate long term growth to endogenous investments in physical, knowledge, and human capital (Romer, 1990; Lucas, 1988). Investments in human capital and new knowledge by firms and HEIs are considered to result in knowledge spillovers, resulting in a positive externality benefiting the local economy, and possibly spilling over to other regional economies. These models allow for the possibility of sustained permanent growth rate differences across regional economies resulting from differences in innovative efforts and capabilities, with new knowledge being subject to increasing returns to scale.

HEIs may be modelled as an input into the knowledge production function (Griliches, 1979 and 1984) which relates innovative outputs, such as patent applications, to innovative inputs such as research and development (R&D) and human capital. Jaffe (1989) analyses the potential importance of geographically based complementarities between university and firm research within the local area, finding that where such complementarities exist, universities are a catalyst for increasing innovation output at the regional level.

Jacobs (1969) argues that knowledge can be divided into two main classifications: codified knowledge and tacit knowledge Codified knowledge is knowledge that has a common interpretation and can be cheaply transferred across agents and space. Conversely, tacit knowledge is costly to transfer across agents and space, requiring proximity (face to face interaction) in order to be absorbed. If much of the newly generated knowledge is tacit, the spillovers will be geographically bounded with benefits decreasing across space. This means that firms closer to the source of the new knowledge will be better able to absorb it, incentivising firms and people to locate in the area. Furthermore, if innovation grows disproportionately with size (Baumol, 2002), then a feedback mechanism between clustering and innovation may occur, similar to the process suggested by Krugman (1991).

Proximity to the primary knowledge source may be insufficient to generate benefits from knowledge production; the region's capabilities to absorb and apply the knowledge may also be critical (Fagerberg, 1987). For example, two regions which increase their local innovative efforts (or that are similarly proximate to new sources of knowledge) may experience significantly different economic growth outcomes if they differ in their ability to extract externally generated knowledge in order to give these ideas economic value. Thus the quality of local human capital may be crucial in generating long term economic benefits from new knowledge. Glaeser et al (1995) examined population growth patterns for over 200 US cities over 1950 to 1990. In testing the importance of a number of initial conditions that included ethnic structure, labour force and educational indicators (plus geographic dummies) the study found that initial education levels of the population were an important

[4] Apatov (2013) provides an in-depth discussion of the relevant literature.

determinant for cities' productivity, positively affecting growth in income, employment, and population.

Duch et al (2011) analysed the channels by which universities contribute to regional growth in Spain (through human capital creation, knowledge generation, and technology transfer). Under all specifications, initial conditions – the share of tertiary educated workforce and the initial stock of patents – were found to have positive and significant growth effects. In contrast, other channels for a university's contribution (university R&D expenditure, R&D incomes, and university internships) were found to be insignificant. Similarly, Trendle et al (2004), applying a spatial lag model to Queensland, found that the proportion of population with a vocational, bachelor or higher degree is an important determinant for local incomes. Wang (2010) found that HEIs contribute to local area growth through their production of skilled graduates, albeit with heterogeneity in effects according to the institution's size, disciplines offered and level of graduates (with business degrees and Masters/Doctoral qualifications having a greater effect). Furthermore, application of a spatial framework showed that such benefits were not limited to the host county, but also positively affected neighbouring counties' employment growth rates. Anderson and Karlson (2005) found that such positive spillover effects extended (in Sweden) to the intra-municipal and intra-regional levels, but not to extra-regional levels, consistent with the localised importance of tacit knowledge.

A common empirical functional specification for the studies cited above is the change-level approach. In this specification, growth rates for the outcome variable of interest (e.g. population, economic activity or incomes) is a function of the levels of pre-existing characteristics (e.g. skills, or stocks of knowledge). While coming from a different theoretical basis, this functional form is essentially the same as that arrived at by Duranton and Turner (2012), and is the functional form that underlies our analysis.

In applying this type of framework, Crescenzi (2005) showed that while R&D investment has a positive and significant effect for a region, innovative efforts will have a better return in regions that are on average more educated and accessible. Sterlacchini (2008) similarly found that local R&D investment was positively associated with economic growth for richer regions but not for poorer regions, whereas an increase in the tertiary educated population share was positive and significant for both types of regions. Rodriguez-Pose and Crescenzi (2008) also find that differences in the education level of the workforce and accessibility to other regions are important factors in translating these investments into economic growth. Thus both distance from the source and skill levels are important complements in gaining benefits from the generation of knowledge.

Mollick and Mora (2012) recognise the potential two-way causation between education levels and growth. To avoid bias, they use a two equation system for growth in population and education level (share of tertiary educated workforce) in the initial period of the analysis. Their study again supports the importance of a tertiary educated workforce for population and employment growth, and note that when estimation does not account for endogeneity, the coefficients understate the importance of education for growth.[5]

These studies together suggest that the presence of HEIs assist local growth, but that a key channel of such influence may be through the production of an educated workforce rather than through the

[5] This finding may imply that HEIs have been explicitly located in otherwise underperforming areas.

direct contribution of an HEI to knowledge production. This latter channel may, however, be dependent on other complementarities such as the relationship with local industry R&D.

In the New Zealand context, Apatov (2013) found that if potential endogeneity in the location of HEIs (universities and polytechnics) is not controlled for, HEIs are found to have a positive link with local population and employment growth. In addition, this growth effect was found to increase in a non-linear manner with increasing levels of population density. However, after controlling for potential endogeneity in HEI location - by instrumenting using population estimates from 60 years earlier - the relationship is insignificant in almost all specifications. In keeping with a number of international studies, however, the share of tertiary qualified working age population in an area is found to be a key driver of economic growth (with and without instrumenting).

2.4 Amenities and Regional Growth

Desmet & Rossi-Hansberg (2013) adapt the Alonso-Mills-Muth model of city structure (Alonso, 1964; Mills, 1967; Muth, 1969; Kulish et al, 2012) to examine the determinants of city size in the US and China. In their theoretical model, an increase in each of city productivity, city efficiency (e.g. of public services) and city amenities leads to an increase in city size. Conceptually, this approach is consistent with the model of Overman et al (2010) in which people migrate between cities to take advantage of higher amenity-adjusted real wages. Desmet & Rossi-Hansberg find strong empirical support for their model, with city amenities playing a particularly important role in determining city size. An important feature of their model is the role played by the retired population. Retirees are found to shift to cities that have high amenities even where those cities are not highly productive.

In considering amenities that affect people’s residential locations, Desmet & Rossi-Hansberg build on prior studies that demonstrate the importance of weather (especially winter and summer temperatures, and precipitation) and coastal locations for determining people’s location decisions within the US (Rappaport, 2007, 2008 and 2009; Rappaport and Sachs, 2003). These studies’ findings regarding the importance of weather for attracting population mirror an earlier finding by Glaeser et al (2001) in this respect. In that study, Glaeser et al describe four critical urban amenities. The first is a rich variety of services and consumer goods including “restaurants, theaters and an attractive mix of social partners”. Larger cities tend to excel in these respects. The second is aesthetics and physical setting, including weather. The third is good public services6 and the fourth is speed or connectivity.7 Each of these factors should therefore be included either directly or indirectly in an empirical model explaining long run population growth.

[6] Good public services may be a positive function both of city size (e.g. a large city is more likely to be able to offer some services, such as a reference library, that cannot be offered in a small town) and of local government efficiency, a factor emphasised by Desmet & Rossi-Hansberg.

[7] With regard to speed, Glaeser et al note that this may be achieved either in a car-based decentralised city or in dense CBD-oriented city.

3. Theory of Population Location

3.1 Population Location Model

Assume that individual i has utility defined over private consumption, C_i (≥0), plus consumption of unpriced amenities available at the core location, S (≥1),[8] where utility derived from the amenities is itself a function of the individual's distance, D_i (≥0), from the core. Greater distance from the amenities leads to lower value being placed on those amenities. For example, a close-up view of a beach may confer greater utility than does a distant view. Similarly, proximity to a social amenity such as a base hospital may confer greater utility through peace of mind (especially for a person prone to illness) than being distant from the hospital. The resulting utility function is given by (1):

$$U_i = C_i^{\beta_i} S^{(\gamma_i + \delta_i D_i)} \quad (1)$$

where β_i>0, δ_i<0, and $\gamma_i+\delta_i D_i$ >0 for any admissible distance, D_i.[9] The individual's budget constraint comprises the individual's earnings, W_i, less expenditure on consumption C_i, (with the consumption price normalised to unity), land L_i, and transport costs T_i. Thus:

$$W_i = C_i + L_i + T_i \quad (2)$$

The wage rate for individual i is set at what the individual could earn at the core location (D_i = 0), w_i, less an individual-specific distance-related discount (at rate q_i) reflecting productivity losses as the individual locates to a more peripheral area. Assuming a linear loss function, we therefore have:

$$W_i = w_i - q_i D_i \quad (3)$$

Land costs are highest at the core location (with L_i = l at D_i = 0) and decline linearly with distance at rate p (which is identical for all individuals). Thus:

$$L_i = l - pD_i \quad (4)$$

Expenditure on transport is an individual-specific increasing function of distance, with transport costs T_i = t_i at D_i = 0, increasing at rate r_i as distance from the core rises. Hence:

$$T_i = t_i + r_i D_i \quad (5)$$

Each of q_i, p, r_i > 0. Substituting (3)-(5) into (2), and denoting $y_i \equiv w_i - l - t_i$ and $z_i \equiv q_i - p + r_i$, yields the budget constraint:

$$y_i = C_i + z_i D_i \quad (6)$$

Maximising (1) subject to (6) gives the solutions for C_i and D_i in (7) and (8):

$$C_i = \frac{z_i \beta_i}{\log(S^{\delta_i})} \quad (7)$$

---

[8] Given the multiplicative utility function, the minimum value for S is 1; at this level (corresponding to zero amenities) the individual's utility is determined solely by private consumption.

[9] Duranton and Turner's (2012) utility function accords distance travelled a positive elasticity, based on an argument that people travel in order to experience amenity services. We consider it more realistic to treat distance travelled as a cost rather than as a benefit.

$$D_i = \frac{y_i}{z_i} - \frac{\beta_i}{\log(S^{\delta_i})} \quad (8)$$

We assume that $z_i<0$, so that individuals are better off financially living distant from the core in order to compensate them for the loss of utility in living away from the core. Taking the parameters as exogenous to the individual, several partial equilibrium results arise in relation to the individual's optimal choices. First, since the denominator of (7) is negative (with S>1), then:

$$\frac{\partial C_i}{\partial z_i} < 0. \text{ Hence: } \frac{\partial C_i}{\partial r_i} < 0, \frac{\partial C_i}{\partial q_i} < 0, \text{ and } \frac{\partial C_i}{\partial p_i} > 0. \text{ In addition: } \frac{\partial C_i}{\partial S_i} < 0.$$

Similarly, from (8) we find:

$$\frac{\partial D_i}{\partial z_i} < 0. \text{ Hence: } \frac{\partial D_i}{\partial r_i} < 0, \frac{\partial D_i}{\partial q_i} < 0, \text{ and } \frac{\partial D_i}{\partial p_i} > 0. \text{ In addition: } \frac{\partial D_i}{\partial S_i} < 0.$$

Thus as the wage premium for living in the core (q_i) increases (and hence z_i rises), people choose to live closer to the core city (D_i falls). Consumption falls as a result of greater expenditure on land. Similarly, as travel costs (r_i) increase, people reduce their distance from the core but at the expense of having to lower consumption. As the price discount for living distant from the core (p_i) increases, people choose to locate further from the core and increase their consumption at the same time. As amenities in the core (S_i) increase, people choose to reduce their distance to the core city but this incurs additional expenditure on land which results in a reduction in their private consumption.

In general equilibrium, land prices (defined by the parameters l_i and p_i) will adjust in response to other factors to effect a spatial equilibrium. For instance, since an increase in amenities within the core causes more people to choose to locate near the core, land prices at the centre will increase and hence l_i will increase. The equilibrium outcome will reflect factors such as land supply elasticities in alternative locations which we do not model here. We assume, however, that planning and topographical constraints are not so rigid as to offset the directions of impact derived from the partial equilibrium results. Table 1 summarises the results from the partial equilibrium model.

Table 1: Summary of Key Partial Equilibrium Model Results

Parameter or Variable	**Effect on Optimal Choices of Individual i**	
	C_i	**D_i**
Core wage premium increase	↓	↓
Land price discount for distance increase	↑	↑
Travel cost increase	↓	↓
Core amenities increase	↓	↓

For the purposes of our analysis, the two most important results in interpreting infrastructure's impact on population location are those for changes in travel cost and core amenities. As core amenities increase, we predict an increase in the number of people choosing to locate closer to the core. To interpret the travel cost result, contrast two cities of equal distance from the core where city A has lower travel costs (r_i) per unit of distance than does city B. In choosing between the two cities, ceteris paribus, individual i will gain greater utility locating in city A than in city B. Thus, holding constant the other parameters faced by each individual, we expect to see greater population in city A than in city B. A reduction in travel costs in city B relative to city A then causes a flow of people to city B, the city that had higher initial travel costs.

3.2 Population Growth Implications of an Amenity Increase

The model in section 3.1 is one of static equilibrium. To convert this model into one that has implications for the determinants of population growth, consider an extension in which total amenities are a function of natural amenities (e.g. climate) and social amenities. Social amenities may include physical amenities (e.g. educational institutions, hospitals, concert halls, etc) and the non-physical "buzz" that comes from being in or near a population agglomeration. Thus, taking the natural amenities as given, S=s(A, P) with $S_A>0$ and $S_P>0$, where A is physical amenities and P is population.[10]

Our initial analysis considered only one urban system. Now compare two urban 'galaxies', j and k, each of which has a core and its own periphery. For convenience, we assume that the peripheries of the two galaxies do not overlap. For individual i to live in galaxy j, it follows that their chosen location within galaxy j, which yields consumption C_i^j and distance from the core D_i^j, with utility U_i^j, is such that $U_i^j \geq U_i^k$ for all possible distances within galaxy k.

Now consider a marginal individual, i, who is initially located optimally within galaxy j, but is almost indifferent to moving to galaxy k. Galaxy k then adds to its amenities, S^k, either through an explicit investment in physical social amenities, A^k, or through an influx of population, P^k, (for instance, as a result of international inward migration to galaxy k). The resulting increase in S^k raises the utility that individual i can gain from locating in galaxy k and, if this increase is sufficient, the individual will switch location from galaxy j to within galaxy k, with utility U_i^k, such that $U_i^k \geq U_i^j$ for all possible distances within galaxy j. This location switch causes a decline in P^j and hence in S^j, while causing a rise in P^k and hence a further rise in S^k. Together, these consequential changes in S^j and S^k will result in more (previously infra-marginal) individuals relocating from galaxy j to galaxy k, with further consequences for amenities in the two locations through the population externalities.

The dynamics resulting from this process will depend on the parameters within the amenities function, *s(A, P)*, as well as on the parameters in the utility function.[11] One possible result is that the entire population in galaxy j relocates over time to galaxy k, consistent with the patterns of urban agglomeration coupled with rural depopulation that has been observed for many decades in developed and emerging market countries (Florida et al, 2008; McCann, 2013).

3.3 Population Growth Implications of a Change in the Cost of Distance

A change in the cost of distance similarly may affect migration from one galaxy to another. Assume that each individual relates to only one core (for instance has a favoured core when visiting department stores, hospitals and concerts), and consider individual i who is located at the outer-most periphery of galaxy j which is adjacent to the outer-most periphery of galaxy k. Now let the

---

[10] The positive partial derivative with respect to P reflects an assumption that congestion effects are outweighed by the positive externality "buzz" effects (although this assumption can be relaxed).

[11] The population relocations will also change the constraint parameters (y_i and z_i) facing the individual in both locations. The nature of these changes will depend on local circumstances (e.g. the nature of topographical constraints on the distance discount) so we do not derive general equilibrium outcomes of the changes.

transport network within galaxy k improve so that its marginal travel costs, r_i^k, decline. From (8), a decline in r_i^k (and hence in z_i^k) results in an increase in distance, D_i^k, for optimal location within galaxy k. This means that individual i who was previously just indifferent between treating the centres of galaxy j and of galaxy k as her core will now be attracted more to galaxy k. Thus the effective population within galaxy k increases and the effective population within galaxy j decreases as a result of the reduction in travel costs within k relative to j. Given our assumption that *S* is an increasing function of *P*, these changes in effective population lead to an improvement in amenities within galaxy k and a decline in amenities within galaxy j, causing migration from galaxy j to galaxy k.

This example also suggests that one galaxy may swallow another as a result of a reduction in transport costs. High transport costs result in the existence of multiple cores and multiple galaxies. Given the increasing returns to scale that arise from *S* being a positive function of *P*, as transport costs reduce, individuals initially located in a small galaxy may all choose to relate primarily to the larger galaxy so accentuating the amenity advantages of that galaxy. Over long periods, therefore, as transport services improve, we expect that the number of core cities will decrease while the size of the remaining core cities will increase.

In the presence of increasing returns to amenities related to population, the effect of a decrease in transport costs on the growth of localities at different distances to the core, is ambiguous. Numerical simulations of our model show that when there are only minor increasing returns of amenities to population, a reduction in transport costs, ceteris paribus, results in a spreading of the population across the relevant galaxy (i.e. a higher average D_i). However, a high rate of agglomeration externalities results in a greater concentration of the population towards the centre of the galaxy (lower average D_i). Furthermore, the ceteris paribus assumption is important here, since other parameters in the constraint functions may respond to the distribution of population. Thus we cannot be definitive on the expected effect of a decrease in transport costs on population growth of localities close to the core relative to those more distant from it. However, acknowledging the ceteris paribus assumption, the more important are increasing returns of amenities to population, the more likely it is that improved transport links will result in a concentration rather than a dispersal of population around the core.

3.4 Model Implications for Population Growth Determinants

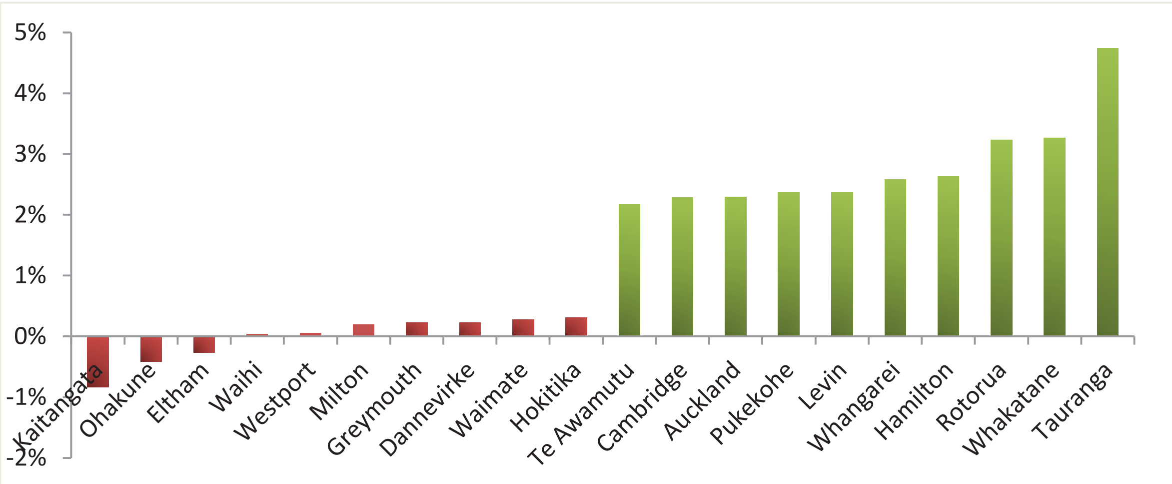
Several empirical implications for the determinants of population growth can be obtained from the model outlined above. First, we hypothesise that an increase in the level of physical amenities (such as educational institutions and hospitals) within a city will lead to an increase in its population growth rate both directly and as an indirect result of the subsequent increase in its population. Second, we expect that an exogenous increase in a city's population (for example, through inward international migration to an internationally connected city) will lead to further sustained population growth. Third, based on results in the literature that agglomeration externalities have increased in recent decades, we anticipate that a reduction of transport costs will increasingly favour growth of cities that are close to the core. In contrast, this relationship may have been reversed in historical periods when agglomeration externalities were less prevalent so that decreased transport costs may historically have favoured growth of localities distant from the core. In testing this last hypothesis, we adopt a maintained assumption (consistent with Glaeser & Kohlhase, 2004) that transport costs have progressively fallen since the start of the twentieth century.

4. Population Growth Empirics

4.1 Modelling Approach

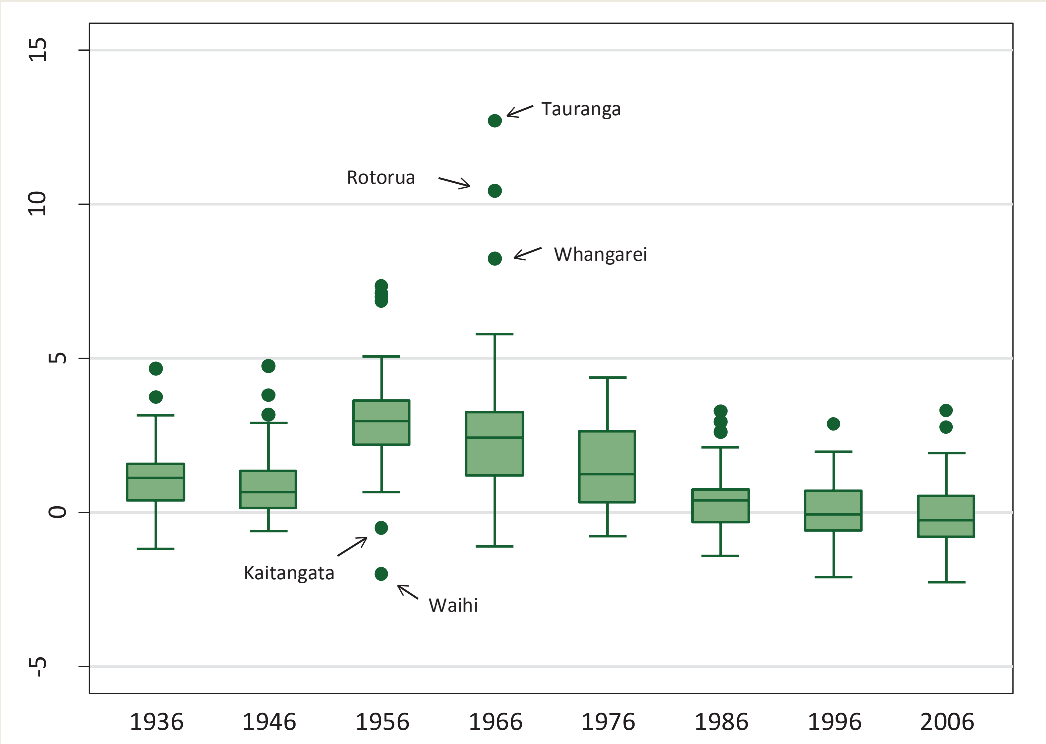
With this theoretical framework in mind, we now examine the historical population growth rates of 56 New Zealand towns over 1926-2006. Our population data consist of eight waves of decennial census figures taken from the New Zealand Urban Population Database, described in detail in Grimes and Tarrant (2013)[12]. The unequal fortunes of New Zealand towns is made plain in Figure 1, which plots the average annual growth rates of the ten fastest- and ten slowest-growing towns. The distribution of urban population growth rates over time is represented via box plots in Figure 2. We see that population growth rates were highest in the first two decades after World War II, and several North Island towns experienced dramatic growth in the decade to 1966. In the two decades between 1986 and 2006, however, slightly over half of the 56 towns experienced negative growth. (Summary statistics for average annual population growth by decade are presented in the Appendix, Table A2).

Figure 1: Annualised Population Growth Rates, Top and Bottom 10 Towns (1926-2006)



[12] Towns are included in the database if they meet at least one of the following criteria: (a) they were categorised as an "urban area" by SNZ in 2006; (b) they were categorised as a "secondary urban area" by SNZ in 1986; (c) the borough population was at least 3,000 in 1956; or (d) the borough population was at least 1,500 in 1926. These criteria ensure that all significant towns in 1926 and 1956 are included, as well as larger urban areas in 1986 and 2006. As detailed in Grimes and Tarrant, the use of 2006 definitions of urban areas and secondary urban areas means that we treat towns that have effectively merged over time as a single urban area (even if they were separate in 1926). Three of the 60 towns in the database were not included in our study as data are not available for some years. Bluff (which may be considered an adjunct town to Invercargill) was found to be an influential negative outlier in some regressions so was also excluded from the analysis, reducing the final number of towns to 56. The data are available here: http://www.motu.org.nz/building-capacity/dataset/new_zealand_urban_population_data

Figure 2: The Distribution of Average Population Growth by Decade



As outlined in the previous section, we expect towns with better wage opportunities and/or better amenities to experience higher population growth, all else being equal. To test this, we estimate variations of the following general model:

$$y_{it} = \sum_j \beta^j x_i^j + \sum_k \phi^k z_{it}^k + \theta D_t + e_{it}$$

where y_{it} is the (geometric) average annual percentage population growth for town i in decade t; each x_i^j represents a fixed wage- or amenity-related characteristic (such as sunshine hours) or an initial condition of interest (e.g. presence of an airport at $t = 0$); each z_{it}^k represents factors that vary over time as well as across towns.; and θD_t is a time fixed effect that picks up the impact of all national demographic and economic factors in decade t. The error term e_{it} is assumed to be correlated within i, and we also test the further assumption of random effects $(e_{it} = u_i + v_{it})$. We estimate the model for the full time span (1926-2006) as well as two subsamples (1926-1966 and 1966-2006) to allow for the possibility that the dynamics of population growth may have changed over time. In particular, some modern infrastructure covariates (such as dummies for airports and polytechnics) are only relevant to the 1966-2006 period, while agglomeration forces may be more relevant to the more modern services-oriented period than in the early period.

4.2 Explanatory Variables

Table 2 categorises the explanatory variables according to whether they are expected to affect population growth through wage opportunities, amenities or both (data sources are detailed in the Appendix, Table A1). Where major fixed infrastructure investments are concerned, we chose long-lived infrastructure that was built at or before the beginning of the time period, when decision makers were unlikely to have had accurate expectations about population growth many decades into the future. Nonetheless, we treat most variables as potentially endogenous.

Variables intended to capture wage prospects include: average land-use capability (LUC), a measure of the suitability of nearby land for agriculture[13]; road distance to port near the start of the time period; dummies for the presence of universities and polytechnics; and a human capital proxy. We do not have longstanding measures of human capital to utilise; instead, we note that throughout post-European settlement of New Zealand, Māori students have consistently had much lower pass rates in school examinations than do Europeans (Pākehā)[14]. Two alternative measures of Māori population as a percentage of total town population are used as proxies for human capital levels, one from 1881 and the other from 1946. For both variables, data were not available for all towns and we had to approximate using the proportion Māori of the nearest neighbour. There are 28 unique values for the 1881 measure but only 13 unique values for the 1946 variable. Although the 1946 measure is much coarser and more likely to be endogenous, it is also more relevant to the time period in question. In light of this trade-off, we choose to use both measures alternatively.

Average annual sunshine hours and rainfall are included as a natural amenities[15], while the presence of an airport could have both amenity and productive value. Similarly, region dummies could capture both productive and amenity differences across regions. We use the following seven regional classifications: *Auckland* (within 200km of Auckland); *Greater Auckland* (all other North Island towns north of Lake Taupo); *Wellington* (within 200km of Wellington); *Greater Wellington* (all other North Island towns south of Lake Taupo); *Christchurch* (within 200km of Christchurch); *Greater Christchurch* (all other towns in Canterbury, Marlborough, Tasman or West Coast regions); and *Dunedin* (Otago and Southland). The regions are thus defined because we expect a town's attractiveness to depend on its distance to the nearest main centre, and on the productivity and amenity benefits of that main centre compared with other regional hubs[16]. In addition to the region dummies, we also include initial road distance from the relevant main centre for towns in the *Auckland, Wellington, Christchurch* and *Dunedin* regions[17]. Once again, initial rather than current road distance is chosen to minimise potential endogeneity.

Finally, we include (log) start of decade population to test for agglomeration externalities. If positive agglomeration externalities outweigh negative effects (e.g. congestion), then larger towns will grow

---

[13] To derive this, we averaged the LUC index values across all 2006 Census meshblocks within each Territorial Local Authority (TLA), weighted by meshblock land area (we also transformed the variable so that higher values corresponded to better agricultural land). Each town was then assigned the average LUC of the TLA that it falls within. A detailed description of the LUC index can be found in Lynn et al. (2009).

[14] For instance, despite improvements in Māori pass-rates in recent decades, Māori pass-rates for NCEA Level 2 in 2012 were 54.2% relative to a non- Māori pass-rate of 74.3% (New Zealand Ministry of Education, 2014). Furthermore, we note the persistence of Māori population proportions over time; the correlation coefficient between the 1946 and 1881 Māori proportions is 0.59.

[15] The Pearson correlation coefficient between sunshine hours and average LUC is -0.06, so we do not interpret sunshine hours as affecting agricultural productivity. We gathered data on hospitals from the 1926 SNZ Yearbook as another amenity measure, but we concluded that the definition of "hospital" at the time was too broad.

[16] We did not distinguish between Southern towns within and beyond 200km of Dunedin because Invercargill is the only town more than 200km away (at a distance of 224km).

[17] We experimented with linear and quadratic distance to each main centre for all towns in the same island (rather than just towns within 200km), and we also tested travel time and the ratio of time to distance. In all cases, there was no significant negative distance effect.

faster than smaller towns. However, omitted variables bias is a potentially serious concern (discussed in the next subsection), and population could also be endogenous if people move to a town in the current period because they expect it to grow in the future (or move away if they expect the town to decline). Our estimation strategy is designed to accommodate such endogeneity concerns.

Aside from the endogeneity issues, our main challenge is multicollinearity amongst the variables of interest. As shown in Table 3, our main infrastructure measures are highly correlated with population, leaving us unable to identify the separate effects of infrastructure variables on population growth. Of course, any variable that has an influence on percentage population growth will eventually be correlated with population level; indeed, the observed correlation between population and distance to port is possibly due to a causal effect on settlement patterns that had already been borne out before the time period that we study. Unfortunately, this means that we don't have the statistical power to separate out any continued effect of proximity to port from the agglomeration effects of population. For this reason, we exclude distance to port from our regressions and bear in mind that the coefficient on start of decade population reflects the influence of omitted town characteristics that are correlated with population levels as well as agglomeration effects.

We face a similar problem with universities, polytechnics and airports. These are investments that were made in towns that were already relatively large, so once again we have little power in testing for their individual effects on population growth. Consequently, our preferred specification excludes universities, polytechnics and airports as well as distance to port, and instead treats lagged population size as a summary variable capturing both amenity and earning effects that are correlated with urban size. Pooled OLS results including the collinear variables are presented in Section 4.3 (Table 6). Note also that there are only six New Zealand cities with a university home campus (four of which had a university before 1926), so we have little variation to work with when trying to single out the effect of universities on growth even in the absence of multicollinearity issues.

Table 2: Explanatory Variables and Reason for Inclusion

Type of Variable:	***Hypothesised to influence:***		
	Wages	***Amenities***	***Both***
Fixed (X_i)	Average land-use capability	Average annual sunshine hours; Average annual rainfall	Region dummies
Initial conditions (X_i)	1932 road distance to port†; Early percentage Maori		Initial road distance to main centre†
1966-2006 only			*Domestic airport in 1966 (dummy)†*
Time-varying (Z_{it})	University at start of decade (dummy)†		Log of start of decade population†
1966-2006 only	*Polytechnic at start of decade (dummy)†*		

Note: † denotes possibly endogenous variables.

Table 3: Correlations with Population

Pearson correlation coefficients	Log of start of decade population		
	1926-2006	**1926-1966**	**1966-2006**
University at start of decade	0.682	0.706	0.685
1932 Road distance to port	-0.468	-0.517	-0.455
Polytechnic at start of decade	-	-	0.679
Airport in 1966	-	-	0.747

4.3 Results

In this section, we present the results from both pooled OLS and random effects models and then discuss the checks we carried out to ensure that the results are robust to potential endogeneity and spatial correlation. We also discuss checks for omitted variables within our preferred equations. Average annual rainfall was included in preliminary regressions (not presented here) and the coefficient was always near zero and insignificant. Omission of the rainfall variable did not affect our other estimates, so it was excluded from subsequent regressions for the sake of parsimony.

Results from the pooled OLS regressions (with clustered standard errors) are shown in Table 4. For each time period, we estimated the regressions using the 1881 and 1946 Maori population measures alternatively. In both cases, the percentage of Maori population is negative and significant in the 1966-2006 time period, but not in the 1926-1966 subsample. This result is consistent with Apatov (2013), which highlights the importance of human capital for regional population growth in recent years.

Average land-use capability has a positive and highly significant relationship with population growth in both the early and latter time periods, suggesting that towns near more productive land have enjoyed long-lasting spill-overs from primary sector profitability. Sunshine hours is another factor with a positive impact on population growth over the whole time period, in line with findings from other countries (see, for example, Rappaport, 2007, 2008 and 2009; Rappaport and Sachs, 2003). This trend could be driven by the migration of retirees or growth of the tourism sector, or it could simply be that sunshine is a luxury good that New Zealanders have increasingly sought out as incomes have grown.

As discussed in Section 3.4, we expect the population level to have a positive effect on growth due to agglomeration externalities and falling transport costs, and the literature suggests that this effect has been stronger in recent decades. Indeed, the coefficient on the population lag is always positive and is significant for the 1966-2006 time period.

We hypothesise that distance to the main centres (particularly Auckland) has a negative impact on growth in the latter period for the same reasons, though it is also possible that distance was beneficial for growth in the early 20th century (if transport costs were prohibitive enough to encourage the growth of distant “mini cores”). The coefficients for all four main centres are positive in the 1926-1966 subsample, with the Dunedin coefficient being significant at the 5% level. In the 1966-2006 period, the coefficients on distance to Auckland, Christchurch and Dunedin are negative but insignificant, while the coefficient on distance to Wellington is positive and significant at the 10% level. These estimated coefficients are broadly consistent with initially positive and subsequently negative distance effects, but are mostly insignificant. The lack of significance is unsurprising for two reasons: the distance variables are only defined for towns within 200km of the relevant main centre, rendering the effective sample size very small (see Appendix Table A3); and these linear distance effects are intended to pick up spatial trends over and above what is picked up by the regional dummies. Note that all regions are penalised relative to the Auckland region in the 1966-2006 subsample, with the effect being most pronounced for Wellington and Greater Wellington regions. Within the Auckland region, distance from Auckland is also penalised, albeit with an insignificant coefficient.

Table 5 presents estimates from a random effects model of population growth with the same covariates, estimated by feasible GLS. The point estimates are approximately the same as those estimated via pooled OLS, but the estimator is more efficient if random effects are present. In all six regressions, the Breusch-Pagan LM test indicates that the random effects model is preferred to the basic pooled OLS model. The random effects results follow the pooled OLS model estimates very closely but for one exception: the estimated coefficient on (log) start of decade population for 1926-2006 is positive and significant in the pooled OLS specification but is close to zero and insignificant in the random effects model. However for the 1966-2006 period, when we expect that agglomeration economies will be more important, the lagged population level variable retains its significant positive effect on the population growth rate.

Finally, pooled OLS results including the collinear infrastructure variables (distance to port and dummies for universities, polytechnics and airports) are shown in Table 6. As expected given our collinearity concerns, these variables are by and large insignificant. The one exception is the dummy for polytechnics in the 1966-2006 subsample, which enters with a negative coefficient that is significant at the 10% level. It is possible that polytechnics have been built in underperforming towns in an attempt to stimulate growth. However, it is more likely that the observed effect is an artefact of the high degree of colllinearity between the variables.

4.4 Test for Omitted Variables

Random effects and pooled OLS models are only valid if there are no omitted variables in the specification. Also, recall that the road distance covariates, region dummies and start of decade population potentially explain population growth through their impact on wages or amenities (or both). In order to check for the possibility of omitted variables in our regressions, and to gain a better understanding of the different factors at work, we turned to the Territorial Local Authority (TLA) rankings created by Donovan (2011). Donovan used Census income and rent data from 1996, 2001 and 2006 to rank TLAs according to their (revealed preference) attractiveness for "business" and "life"[18]. We took the average rankings across the three years as measures of the value accorded to wages and amenities respectively towards the end of our sample. We added a quadratic in each ranking to our population growth regressions (see Appendix Tables A4 and A5) to test whether

[18] Donovan calculated the *life* index as $(r - w)$, where r is the average rent paid by households in the TLA adjusted for housing quality (number of rooms, etc.), and w is the average household income in the TLA, adjusted for observable characteristics such as education level and household size. This index reflects a spatial equilibrium approach in which people pay high rents relative to wages so as to access positive local amenities. The *business* index is defined as $(r + w)$, with household rent proxying for commercial rent. This index also reflects a spatial equilibrium approach in which firms that choose a highly productive locality can pay higher wages and must pay higher rents to reflect the more productive location.

amenity or productive factors that are reflected in rents and wages added significantly to our included explanatory variables in explaining population growth. The coefficients on the quadratics were insignificant in all cases, indicating that there is no evidence that our model omits important factors affecting rents and wages that are not already picked up by our existing covariates.

Separately, we regressed each of the business and amenity ranking variables against the 2006 values of our covariates to analyse which of our covariates influence (revealed preference) amenity and productivity values across towns at the end of our sample. The results are shown in Table 7 (noting that negative coefficients correspond to higher rankings). Land-use capability is reflected significantly in business rankings with the expected sign. Sunshine hours has the expected sign for amenity rankings but is not significant. The population level strongly predicts both the amenity and business rankings, as expected. Finally, note that all regions fare better than the Auckland region for amenities but worse for “business”, implying that the Auckland region’s higher population growth over 1966-2006 (once other factors have been controlled for) was driven principally by wage considerations rather than quality of life.

Having tested for the robustness of our results to potential omitted variables, we proceed to test the robustness of our results to potential endogeneity and spatial correlation. Since the 1881 and 1946 Maori population measures yield the same qualitative results for both the pooled and random effects models, we adopt the 1881 measure from now on as it is more clearly exogenous than the 1946 measure.

TABLE 4:	*Dependent variable: Decade average annual % population growth*					
Pooled OLS estimates	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	1926-2006		**1926-1966**		**1966-2006**	
Average land-use capability	0.275***	0.268***	0.344**	0.341**	0.209***	0.192***
	(0.096)	(0.098)	(0.148)	(0.155)	(0.071)	(0.069)
Average annual sunshine hours (00s)	0.262***	0.267***	0.278***	0.273***	0.239***	0.252***
	(0.043)	(0.043)	(0.059)	(0.061)	(0.038)	(0.037)
Initial road miles to Auckland (00s)	-0.221	-0.194	0.233	0.170	-0.862	-0.768
	(0.615)	(0.621)	(0.802)	(0.803)	(0.698)	(0.703)
Initial road miles to Wellington (00s)	0.353	0.401	0.111	0.079	0.672	0.814*
	(0.499)	(0.457)	(0.628)	(0.664)	(0.546)	(0.445)
Initial road miles to Christchurch (00s)	-0.150	-0.138	0.035	0.016	-0.296	-0.252
	(0.420)	(0.431)	(0.351)	(0.345)	(0.858)	(0.902)
Initial road miles to Dunedin (00s)	0.207	0.216	0.610**	0.607**	-0.241	-0.216
	(0.149)	(0.151)	(0.297)	(0.294)	(0.217)	(0.207)
Region (omitted category: Auckland)						
Wellington	-0.677	-0.711	0.196	0.146	-1.700**	-1.770***
	(0.627)	(0.600)	(0.827)	(0.864)	(0.692)	(0.642)
Christchurch	-0.434	-0.458	-0.248	-0.318	-0.766	-0.796
	(0.562)	(0.556)	(0.613)	(0.645)	(0.844)	(0.856)
Greater Auckland	0.975	0.914	2.167***	2.271***	-0.343	-0.587
	(0.598)	(0.606)	(0.769)	(0.779)	(0.668)	(0.703)
Greater Wellington	-0.978*	-0.924*	-0.591	-0.673	-1.460**	-1.288*
	(0.536)	(0.548)	(0.628)	(0.631)	(0.663)	(0.692)
Greater Christchurch	-0.584	-0.596	-0.002	-0.104	-1.237*	-1.218*
	(0.561)	(0.552)	(0.694)	(0.724)	(0.700)	(0.684)
Dunedin	-0.651	-0.652	-0.212	-0.314	-1.173	-1.127
	(0.602)	(0.601)	(0.777)	(0.797)	(0.710)	(0.712)
Percentage Maori 1881	-0.003		0.006		-0.011***	
	(0.003)		(0.006)		(0.003)	
Percentage Maori 1946		-0.054		0.049		-0.164**
		(0.070)		(0.136)		(0.074)
Log of start of decade population	0.121*	0.128**	0.137	0.124	0.151**	0.174***
	(0.061)	(0.061)	(0.096)	(0.095)	(0.058)	(0.062)
N	448	448	224	224	224	224
R-squared	0.597	0.597	0.527	0.526	0.631	0.623

All regressions include an intercept and time fixed effects. Standard errors clustered on town in parentheses.

* p<.1, ** p<.05, *** p<.01

TABLE 5: **Random effects model, feasible GLS estimates**	***Dependent variable: Decade average annual % population growth***					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	1926-2006		**1926-1966**		**1966-2006**	
Average land-use capability	0.288***	0.287***	0.342**	0.338**	0.217***	0.202***
	(0.098)	(0.100)	(0.148)	(0.155)	(0.071)	(0.070)
Average annual sunshine hours (00s)	0.279***	0.283***	0.275***	0.271***	0.246***	0.259***
	(0.045)	(0.046)	(0.059)	(0.061)	(0.039)	(0.038)
Initial road miles to Auckland (00s)	-0.417	-0.378	0.280	0.215	-0.961	-0.872
	(0.642)	(0.657)	(0.801)	(0.802)	(0.675)	(0.682)
Initial road miles to Wellington (00s)	0.152	0.179	0.161	0.136	0.603	0.732
	(0.525)	(0.486)	(0.643)	(0.679)	(0.562)	(0.466)
Initial road miles to Christchurch (00s)	-0.272	-0.260	0.065	0.046	-0.343	-0.302
	(0.278)	(0.288)	(0.371)	(0.363)	(0.801)	(0.839)
Initial road miles to Dunedin (00s)	0.219	0.223	0.609*	0.607*	-0.236	-0.213
	(0.233)	(0.227)	(0.320)	(0.315)	(0.248)	(0.233)
Region (omitted category: Auckland)						
Wellington	-0.674	-0.656	0.194	0.131	-1.722***	-1.777***
	(0.606)	(0.596)	(0.840)	(0.878)	(0.668)	(0.627)
Christchurch	-0.496	-0.465	-0.234	-0.318	-0.809	-0.826
	(0.492)	(0.507)	(0.623)	(0.655)	(0.794)	(0.809)
Greater Auckland	0.834	0.757	2.209***	2.318***	-0.402	-0.658
	(0.582)	(0.594)	(0.792)	(0.805)	(0.647)	(0.682)
Greater Wellington	-1.150**	-1.091**	-0.552	-0.636	-1.544**	-1.380**
	(0.517)	(0.540)	(0.639)	(0.645)	(0.636)	(0.669)
Greater Christchurch	-0.769	-0.718	0.040	-0.077	-1.321**	-1.290*
	(0.532)	(0.549)	(0.709)	(0.738)	(0.673)	(0.665)
Dunedin	-0.797	-0.743	-0.178	-0.293	-1.244*	-1.188*
	(0.591)	(0.608)	(0.794)	(0.813)	(0.685)	(0.693)
Percentage Maori 1881	-0.004		0.007		-0.012***	
	(0.003)		(0.006)		(0.003)	
Percentage Maori 1946		-0.041		0.046		-0.160**
		(0.070)		(0.138)		(0.078)
Log of start of decade population	0.023	0.031	0.162	0.148	0.119**	0.140**
	(0.065)	(0.067)	(0.102)	(0.102)	(0.058)	(0.061)
N	448	448	224	224	224	224
R-squared (overall)	0.593	0.593	0.527	0.526	0.630	0.622
Rho	0.194	0.194	0.256	0.258	0.285	0.302
P-value: Breusch-Pagan LM test for random effects	0.000***	0.000***	0.002***	0.002***	0.007***	0.003***
Corr(log population, random effect)	0.253	0.255	-0.049	-0.050	0.104	0.107

All regressions include an intercept and time fixed effects. Standard errors clustered on town in parentheses.

* p<.1, ** p<.05, *** p<.01

TABLE 6: ***Pooled OLS estimates including collinear variables***	***Dependent variable: Decade average annual % population growth***					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	1926-2006		**1926-1966**		**1966-2006**	
Average land-use capability	0.281***	0.277***	0.330**	0.328**	0.244***	0.227***
	(0.100)	(0.103)	(0.152)	(0.161)	(0.083)	(0.085)
Average annual sunshine hours (00s)	0.244***	0.248***	0.269***	0.266***	0.191***	0.203***
	(0.044)	(0.046)	(0.063)	(0.064)	(0.048)	(0.050)
Initial road miles to Auckland (00s)	-0.306	-0.262	-0.010	-0.074	-0.945	-0.807
	(0.579)	(0.591)	(0.778)	(0.781)	(0.604)	(0.635)
Initial road miles to Wellington (00s)	0.223	0.270	-0.126	-0.162	0.447	0.594
	(0.528)	(0.499)	(0.614)	(0.640)	(0.586)	(0.511)
Initial road miles to Christchurch (00s)	-0.378	-0.356	-0.322	-0.344	-0.741	-0.665
	(0.456)	(0.465)	(0.453)	(0.449)	(0.925)	(0.968)
Initial road miles to Dunedin (00s)	-0.039	-0.013	0.159	0.140	-0.676*	-0.583
	(0.293)	(0.295)	(0.418)	(0.423)	(0.401)	(0.398)
Region (omitted category: Auckland)						
Wellington	-0.628	-0.641	0.236	0.200	-1.583***	-1.622***
	(0.551)	(0.545)	(0.746)	(0.774)	(0.544)	(0.522)
Christchurch	-0.394	-0.399	-0.181	-0.230	-0.654	-0.663
	(0.476)	(0.486)	(0.557)	(0.597)	(0.687)	(0.722)
Greater Auckland	0.880	0.829	1.945**	2.018**	-0.497	-0.683
	(0.592)	(0.604)	(0.768)	(0.764)	(0.612)	(0.683)
Greater Wellington	-1.089**	-1.024*	-0.835	-0.913	-1.583***	-1.401**
	(0.520)	(0.537)	(0.616)	(0.614)	(0.589)	(0.654)
Greater Christchurch	-0.699	-0.679	-0.275	-0.360	-1.378**	-1.323**
	(0.546)	(0.538)	(0.674)	(0.711)	(0.636)	(0.650)
Dunedin	-0.623	-0.611	-0.090	-0.157	-1.102*	-1.069*
	(0.559)	(0.566)	(0.763)	(0.792)	(0.600)	(0.619)
Percentage Maori 1881	-0.003		0.005		-0.011***	
	(0.003)		(0.006)		(0.003)	
Percentage Maori 1946		-0.049		0.044		-0.147**
		(0.073)		(0.141)		(0.063)
Log of Start of Decade Population	0.184*	0.185*	0.286*	0.282*	0.230**	0.221**
	(0.103)	(0.105)	(0.147)	(0.146)	(0.106)	(0.108)
Distance to port 1932	-0.020	-0.040	0.224	0.252	-0.118	-0.202
	(0.324)	(0.331)	(0.483)	(0.487)	(0.265)	(0.270)
University at start of decade	-0.463	-0.428	-0.912	-0.944	-0.553	-0.405
	(0.464)	(0.461)	(0.622)	(0.627)	(0.475)	(0.474)
Polytechnic at start of decade					-0.244*	-0.269*
					(0.131)	(0.137)
Airport 1966					0.156	0.185
					(0.219)	(0.220)
N	448	448	224	224	224	224
R-squared	0.599	0.598	0.533	0.532	0.641	0.632

All regressions include an intercept and time fixed effects. Standard errors clustered on town in parentheses.

* p<.1, ** p<.05, *** p<.01

TABLE 7: *TLA rankings as dependent variable, OLS estimates*	**(1) Amenity ranking**	**(2) Business ranking**	**(continued)**	**(1) Amenity ranking**	**(2) Business ranking**
Average land-use capability	1.738	-4.502**	*Region (omitted category: Auckland)*		
	(2.387)	(1.956)	Wellington	-10.069	13.557
Average annual sunshine hours (00s)	-1.246	-0.744		(14.625)	(11.245)
	(1.484)	(1.109)	Christchurch	-18.785	16.945*
Initial road miles to Auckland (00s)	-6.200	13.902		(14.443)	(9.570)
	(13.097)	(8.641)	Greater Auckland	-6.424	17.368*
Initial road miles to Wellington (00s)	-7.341	18.700*		(14.264)	(9.009)
	(7.827)	(9.745)	Greater Wellington	-5.107	24.424**
Initial road miles to Christchurch (00s)	13.091**	12.637***		(13.440)	(9.242)
	(6.401)	(4.607)	Greater Christchurch	-26.253	24.754*
Initial road miles to Dunedin (00s)	32.213***	-10.761		(17.549)	(13.098)
	(6.361)	(9.914)	Dunedin	-43.703***	45.594***
Log of 1996 Population	-4.102***	-4.261***		(14.919)	(11.679)
	(1.354)	(1.170)	N	56	56
Percentage Maori 1881	0.170	0.007	R-squared	0.537	0.722
	(0.119)	(0.100)	Intercept	Y	Y

*Robust standard errors in parentheses. * p<.1, ** p<.05, *** p<.01*

4.5 Endogeneity

Most of the significant variables in our regressions are clearly exogenous (land-use capability, sunshine hours, 1881 percentage Maori, and the regional categories). Initial road distance to each main centre could be endogenous if faster-growing towns enjoy more direct roads as a result of their growth[19]. Also, road distance is significant only for Dunedin over 1926-1966, and the coefficient is *positive*, so any reverse causation would understate rather than exaggerate the effect of distance. Nonetheless, we derive two-stage least squares (2SLS) estimates of the road distance effects as a robustness check. We instrument for road distance to each main centre using the straight-line distance and the distance by rail in 1909, the year in which the North Island main trunk was completed (though the rail routes were decided well before 1909). We expect rail distance to pick up exogenous influences on road distance such as topography.

The estimated positive effect of log population could also reflect endogeneity bias if population at the start of the decade includes a significant number of people who migrated to the town in expectation of future growth (or equally, the absence of people who moved away in expectation of decline). We generate 2SLS estimates as a precaution; however it is worth noting that this potential endogeneity is in a way self-moderating, since any population influx results in a ceteris paribus decrease in the next period's percentage population growth rate. Ideally, we would use a long lag of population as an instrument, but complete population data for all towns is not available prior to 1926. We use the 40-year lag of log population in the 1966-2006 model, but resort to a number of other instruments for the 1926-2006 and 1926-1966 regressions. These instruments include:

[19] It is also conceivable that underperforming towns are targeted for new roads, but we consider this unlikely.

dummies for the status of rail in 1880 ("has rail", "priority for rail" and "not priority", as determined by the 1880 Railway Commission Report); rainfall and temperature variables; dummies for dairy factories and meatworks in 1920; ports and port tonnage in 1906; dummies for railroads, coach routes and steamer ports in 1901; population in 1901 and a dummy for missing 1901 population data; and a dummy for presence of a university in 1901 (consult Appendix Table A1 for a full list of data sources).

Despite the range of historical data at our fingertips, we do not possess strong instruments for population over the first half of our sample period. This is evidenced by the Angrist-Pischke first-stage F statistics of 21.4 and 22.0 for the 1926-2006 and 1926-1966 regressions, compared with 719.2 in the 1966-2006 regression (where we are able to make use of the 40-year lag)[20]. However, we note that the log population coefficient is only consistently significant in the 1966-2006 pooled and random effects regressions, the time period for which we are not hampered by a weak instrument problem.

The 2SLS estimates, reported in Table 9, indicate that our main results are not driven by endogeneity[21]. We see that the coefficient on distance to Dunedin in the 1926-1966 period remains positive and significant, and is in fact slightly larger than the OLS and random effects estimates. Meanwhile, the coefficients for the log population lag are remarkably similar to our OLS results and their statistical significance is unchanged. These findings are corroborated by the results of Sargan-type regressor endogeneity tests, displayed in Table 8[22]. In all cases, we fail to reject the null hypothesis of exogeneity.

TABLE 8:		**P-values**	
Regressor Endogeneity Tests	**1926-2006**	**1926-1966**	**1966-2006**
(H_0: Regressor can be treated as exogenous)			
Initial road distance to Auckland	0.460	0.380	0.217
Initial road distance to Wellington	0.293	0.950	0.470
Initial road distance to Christchurch	0.932	0.708	0.965
Initial road distance to Dunedin	0.227	0.214	0.264
Log of start of decade population	0.135	0.874	0.955

[20] All 2SLS estimates and diagnostics were computed in Stata using the –ivreg2- command written by Baum, Schaffer and Stillman (2010).
[21] Table 9 also reports p-values from the Hansen (robust) over-ID test as well as Kleibergen-Paap F-statistics. The Kleibergen-Paap F-statistic is an overall indicator of instrument strength (as opposed to the Angrist-Pischke F-statistic, which indicates the strength of instruments for individual endogenous regressors).
[22] Details of the endogeneity tests are discussed in Baum, Schaffer and Stillman (2003).

TABLE 9: *2SLS estimates*	*Dependent variable: Decade average annual % population growth* **1926-2006**	**1926-1966**	**1966-2006**
Average land-use capability	0.272***	0.346**	0.209***
	(0.094)	(0.148)	(0.070)
Average annual sunshine hours	0.257***	0.281***	0.238***
	(0.041)	(0.059)	(0.038)
Initial road distance to Auckland	-0.036	0.406	-0.804
	(0.642)	(0.807)	(0.703)
Initial road distance to Wellington	0.415	0.102	0.658
	(0.507)	(0.598)	(0.554)
Initial road distance to Christchurch	-0.117	0.021	-0.304
	(0.472)	(0.340)	(0.862)
Initial road distance to Dunedin	0.255	0.734**	-0.227
	(0.156)	(0.296)	(0.210)
Region (omitted category: Auckland)			
Wellington	-0.565	0.364	-1.644**
	(0.671)	(0.878)	(0.690)
Christchurch	-0.298	-0.072	-0.715
	(0.630)	(0.683)	(0.847)
Greater Auckland	1.125*	2.311***	-0.297
	(0.660)	(0.832)	(0.667)
Greater Wellington	-0.815	-0.436	-1.413**
	(0.612)	(0.722)	(0.664)
Greater Christchurch	-0.414	0.152	-1.189*
	(0.657)	(0.823)	(0.700)
Dunedin	-0.527	-0.126	-1.136
	(0.676)	(0.860)	(0.712)
Percentage Maori 1881	-0.002	0.006	-0.011***
	(0.003)	(0.006)	(0.003)
Log of Start of Decade Population	0.149*	0.111	0.151**
	(0.077)	(0.104)	(0.057)
N	448	224	224
R-squared	0.596	0.526	0.631
Hansen's J p-values	0.091*	0.203	0.461
Kleibergen-Paap F-statistic	17.366	16.883	393.391

All regressions include an intercept and time fixed effects. Standard errors clustered on town in parentheses.
* p<.1, ** p<.05, *** p<.01

4.6 Spatial correlation

All our specifications to date have modelled the interrelations between each town and its nearest neighbours explicitly, by including regional controls and distance effects for towns within 200km of a main centre. However, it is possible that our model fails to adequately capture more complex spatial interactions that may be at work.

To explore this possibility, we test the significance of two different spatial augmentations to our model. The first of these is the spatial autoregressive model (SAR), which assumes that a town's population growth is influenced by a weighted average of neighbouring towns' population growth levels in the same period (i.e. a spatial lag term). The second augmentation is the spatial error model (SEM), which assumes that population shocks flow through nearby towns (i.e. a spatial error term). We defined the weights used to calculate the spatial lag and spatial error terms as equal to the inverse of the distance between each town, and set the weight to zero for towns not in the same island (i.e. we assumed that North Island towns exert no influence on South Island towns, and vice versa).

We estimated both the SAR and SEM models within our preferred random effects framework (as indicated by the Breusch-Pagan LM test) and the results are shown in Table 10.[23] Both the spatial lag and spatial error terms are insignificant in each time period and our existing coefficient estimates are unchanged, leading us to conclude that our original model does not suffer from spatial misspecification.

[23] Spatial random effects models are implemented using the –xsmle- Stata command introduced by Belotti et al. (2013).

TABLE 10: *Spatial random effects models, maximum likelihood estimates*	*Dependent variable: Decade average annual % population growth*					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	1926-2006		**1926-1966**		**1966-2006**	
	SAR	**SEM**	**SAR**	**SEM**	**SAR**	**SEM**
Rho (spatial lag term)	**0.105**		**0.027**		**0.046**	
	(0.112)		**(0.156)**		**(0.094)**	
Lambda (spatial error term)		**0.142**		**0.110**		**0.045**
		(0.110)		**(0.162)**		**(0.120)**
Average land-use capability	0.282***	0.282***	0.342**	0.341**	0.213***	0.213***
	(0.094)	(0.095)	(0.142)	(0.143)	(0.069)	(0.069)
Average annual sunshine hours	0.270***	0.273***	0.275***	0.277***	0.240***	0.242***
	(0.042)	(0.042)	(0.058)	(0.057)	(0.037)	(0.037)
Initial road distance to Auckland	-0.355	-0.357	0.256	0.249	-0.905	-0.904
	(0.621)	(0.621)	(0.771)	(0.769)	(0.666)	(0.669)
Initial road distance to Wellington	0.219	0.210	0.140	0.134	0.640	0.638
	(0.512)	(0.516)	(0.607)	(0.608)	(0.535)	(0.540)
Initial road distance to Christchurch	-0.207	-0.228	0.055	0.047	-0.299	-0.310
	(0.323)	(0.320)	(0.345)	(0.341)	(0.811)	(0.812)
Initial road distance to Dunedin	0.205	0.210	0.607**	0.606**	-0.243	-0.239
	(0.189)	(0.199)	(0.294)	(0.287)	(0.220)	(0.221)
Region (omitted category: Auckland)						
Wellington	-0.638	-0.678	0.203	0.183	-1.690**	-1.707***
	(0.600)	(0.598)	(0.808)	(0.806)	(0.660)	(0.660)
Christchurch	-0.424	-0.484	-0.219	-0.250	-0.777	-0.787
	(0.503)	(0.502)	(0.612)	(0.595)	(0.795)	(0.801)
Greater Auckland	0.893	0.873	2.191***	2.176***	-0.360	-0.369
	(0.585)	(0.585)	(0.749)	(0.749)	(0.637)	(0.641)
Greater Wellington	-1.049**	-1.100**	-0.559	-0.586	-1.477**	-1.495**
	(0.517)	(0.517)	(0.613)	(0.610)	(0.634)	(0.635)
Greater Christchurch	-0.631	-0.722	0.045	-0.003	-1.248*	-1.273*
	(0.544)	(0.535)	(0.695)	(0.674)	(0.671)	(0.669)
Dunedin	-0.645	-0.749	-0.165	-0.204	-1.162*	-1.202*
	(0.595)	(0.590)	(0.737)	(0.765)	(0.682)	(0.680)
Percentage Maori 1881	-0.004	-0.004	0.006	0.007	-0.011***	-0.011***
	(0.003)	(0.003)	(0.006)	(0.005)	(0.003)	(0.003)
Log of Start of Decade Population	0.056	0.055	0.150	0.148	0.137**	0.137**
	(0.060)	(0.059)	(0.095)	(0.093)	(0.057)	(0.058)
N	448	448	224	224	224	224
R-squared	0.595	0.595	0.527	0.527	0.632	0.631

All regressions include an intercept and time fixed effects. Standard errors clustered on town in parentheses.

* p<.1, ** p<.05, *** p<.01

5. Conclusions

We have analysed the key growth determinants of 56 New Zealand towns and cities over eight decades to 2006. The analysis has been made possible through access to long run city population data and data for many other variables from Statistics New Zealand Yearbooks and other official data sources dating back to the start of the twentieth century.

Using a revealed preference framework, we argue that urban areas grow if they have a desirable combination of amenities and real earning opportunities relative to alternative locations. This framework is formalised within a theoretical model that includes distance-related and amenity effects on individual utility, incomes and costs. A number of factors may contribute to earnings opportunities and/or amenities including transport links, social infrastructure, benefits of location in a large population area, and natural amenities.

In testing our model, we face a number of econometric issues. First, there is a strong positive correlation of urban population with of many of our infrastructure variables. This makes it difficult to identify urban growth impacts, for instance of higher educational institutions, that are separate from their location within a larger urban area. We deal with this issue by using lagged population size as a summary variable capturing both amenity and earning effects that are correlated with urban size. Second, we are cognisant that a number of transport and social amenity variables that we hypothesise are important determinants of urban growth may be endogenously determined. We compile a range of exogenous or long pre-determined variables to use as instruments and test for variable exogeneity using over-ID tests. This testing is made possible through access to the long history of official data in New Zealand. The over-ID tests confirm that all variables in our core regression can be treated as exogenous. Third, we recognise that spatial lag or spatial error processes may affect urban growth patterns. We test whether such processes are important in explaining urban growth over the period, finding little empirical support for their presence once other spatial variables are controlled for explicitly.

Following these tests, we adopt a random effects model as our preferred econometric specification. Based on this specification, we find that five dominant factors have impacted positively on urban growth, especially since 1966: local land use capability, sunshine hours, human capital, population size and proximity to the country’s dominant city, Auckland.

Each of the last three of these elements is potentially a source of policy intervention. First, human capital can be raised through a generalised increase in the national standard of human capital and, at the local level, can be raised by developing and attracting high human capital to the area. The presence of universities (and possibly other HEIs) is correlated with an urban area having high relative human capital, although the causality in this relationship is difficult to establish. Second, proximity to Auckland can be improved through the upgrading of transport links that make it easier for firms and people to locate near to, but outside of, Auckland while still accessing some of the amenity and productivity benefits offered by the city. Third, the importance of population size can be interpreted in an international context. Even Auckland, New Zealand’s largest city, is small by international comparisons and is only the fifth largest urban area in Australasia. To the extent that urban growth across Australasia is determined by similar factors to urban growth within New Zealand, there is a case that policy should at least facilitate, and certainly not overly constrain, the size of New Zealand's largest population centre; otherwise the risk is that growth will increasingly be located in Australia's four largest cities rather than in Auckland and its surrounding regions.

References

Alexander Turnbill Library (2006). *Map New Zealand.* Auckland: Random House.

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

Anderson, M., & Karlson, C. (2007). Knowledge in regional economic growth - the role of knowledge accessibility. *Industry and Innovation, 14(2)*, 129-149.

Apatov, E. (2013). *Higher Education Institutions and Economic Growth within New Zealand*. Master of Commerce Dissertation, University of Auckland.

Aschauer, D.A., "Is Public Expenditure Productive?," *Journal of Monetary Economics* 23 (1989), 177-200.

Baum, C. F., Schaffer, M. E., & Stillman, S. (2003). Instrumental variables and GMM: Estimation and testing. *Stata Journal*, *3*(1), 1-31.

Baum, C.F., Schaffer, M.E., Stillman, S. (2010). ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression. Retrieved from http://ideas.repec.org/c/boc/bocode/s425401.html

Baumol, W. (2002). *The free-market innovation machine: Analysing the growth miracle of capitalism*. Princeton, NJ: Princeton University Press.

Baum-Snow, N., "Changes in Transportation Infrastructure and Commuting Patterns in US Metropolitan Areas, 1960–2000," *American Economic Review* 100 (2010), 378-382.

Behrens, K. and P.M. Picard, "Transportation, freight rates, and economic geography," *Journal of International Economics* 85 (2011), 280-291.

Belotti, F., Hughes, G., & Mortari, A. P. (2013). XSMLE: Stata module for spatial panel data models estimation. *Statistical Software Components*.

Blum, U., "Effects Of Transportation Investments On Regional Growth: A Theoretical And Empirical Investigation," *Papers in Regional Science* 49 (1982), 169-184.

Coleman, A. 2012. "The Effect of Transport Infrastructure on Home Production Activity: Evidence from Rural New York, 1825-1845",Working Paper 12-01, Wellington: Motu Economic and Public Policy Research.

Crescenzi, R. (2005). Innovation and regional growth in the enlarged Europe: The role of local innovative capabilities, peripherally and education. *Growth and Change, 36(4)*, 471-507.

Desmet, K. & Rossi-Hansberg, E. 2013. Urban Accounting and Welfare, *American Economic Review*, 103(6), 2296-2327.

Ding, C., "Transport Development, Regional Concentration and Economic Growth," *Urban Studies* 50 (2013), 312-328.

Donovan, S. (2011). *Space matters: Agglomeration economies and spatial competition in New Zealand* (Unpublished master's thesis). VU University, Amsterdam, the Netherlands.

Donovan, S. and I. Munro, "Impact of urban form on transport and economic outcomes," (Wellington, New Zealand: New Zealand Transport Agency, 2013).

Dougherty, I. (1999). *Bricklayers and Mortarboards: A History of New Zealand Polytechnics and Institutes of Technology*, Palmerston North: Dunmore Press.

Duch, N., Garcia-Estevez, J., & Parellada, M. (2011). *Universities and regional economic growth in Spanish regions* (Unpublished research report). Barcelona Institute of Economics, Barcelona, Spain.

Duflo, E., Pande, R., 2007. Dams. *Quarterly Journal of Economics* 122(2), 601–646.

Duranton, G., P. Morrow and M. Turner, "Roads and Trade: Evidence from the U.S," University of Toronto, Department of Economics, Working Papers, 2013), Unknown.

Duranton, G. & Turner, M.A. (2012). "Urban Growth and Transportation," *Review of Economic Studies* 79, 1407-1440.

Fagerberg, J. (1987). A technology gap approach to why growth rates defer. *Research policy, 16(2-4)*, 87-99.

Fernald, J.G., "Roads to Prosperity? Assessing the Link between Public Capital and Productivity," *American Economic Review* 89 (1999), 619-638.

Florida, R., T. Gulden and C. Mellander. 2008. “The Rise of the Mega-Region.” *Cambridge Journal of Regions, Economy and Society*, 1(3), 459-476.

Fujita, M. and J.-F.o. Thisse, *Economics of agglomeration: cities, industrial location, and regional growth* (New York: Cambridge University Press, 2002)

Gibbons, S., Lyytikainen, T., Overman, H., Sanchis-Guarner, R., 2012. New Road Infrastructure: The Effects on Firms. SERC Discussion Paper 117. London: Spatial Economics Research Centre.

Glaeser, E.L., "Are Cities Dying?," *Journal of Economic Perspectives* 12 (1998), 139-160.

Glaeser, E. & J. Gottlieb. 2009. “The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States”, *Journal of Economic Literature* 47(4), 983-1028.

Glaeser, E.L. and J.E. Kohlhase (2004), “Cities, Regions and the Decline of Transport Costs”, *Papers in Regional Science* 83, 197-228.

Glaeser, Edward L., Jed Kolko, and Albert Saiz. 2001. “Consumer City.” *Journal of Economic Geography*, 1 (1): 27–50.

Glaeser, E. L., Scheinkman, J. A., & Shleifer, A. (1995). Economic growth in a cross-section of cities. *Journal of Monetary Economics, 36(1)*, 117-143.

GPS Visualizer (2014). Straight-line distances and spatial weights calculated using webpage *GPS Visualizer: Great Circle Distance Maps, Airport Routes, & Degrees/Minutes/Seconds Calculator,* *http://www.gpsvisualizer.com/calculators*

Graham, D.J., S. Gibbons and R. Martin, "Transport Investment and the Distance Decay of Agglomeration Benefits," (London, United Kingdom: Imperial College of London, 2009).

Griliches, Z. (1979). Issues in assessing the contribution of R&D to productivity growth. *Bell Journal of Economics, 10(1)*, 92-116.

Griliches, Z. (1984). *R&D, patents, and productivity*. Chicago, IL: University of Chicago Press.

Grimes, A. 2014. Infrastructure and Regional Economic Growth. Chapter 28 in: *Handbook of Regional Science* (eds. M. Fischer & P. Nijkamp). Heidelberg: Springer.

Grimes, A. and Y. Liang, "Bridge to Somewhere: Valuing Auckland's Northern Motorway Extensions," *Journal of Transport Economics & Policy* 44 (2010), 287-315.

Grimes, Arthur & Nicholas Tarrant. 2013. "An Urban Population Database", Motu Working Paper 13-07, Wellington: Motu.

Grimes, A. and C. Young, "Spatial effects of urban rail upgrades," *Journal of Transport Geography* 30 (6// 2013), 1-6.

Guellec, D., & Loannidis, E. (1997). Causes of fluctuations in R&D expenditures: A quantitative analysis. *OECD Journal: Economic Studies, 29(2)*, 124-138.

Haughwout, A.F., "Public Infrastructure Investments, Productivity and Welfare in Fixed Geographic Areas," *Journal of Public Economics* 83 (2002), 405-428.

Holtz-Eakin, D. and A.E. Schwartz, "Spatial Productivity Spillovers from Public Infrastructure: Evidence from State Highways," *International Tax and Public Finance* 2 (1995), 459-468.

Jacobs, J. (1969). *The Economy of Cities*. New York City, NY: Random House.

Jaffe, A. B. (1989). Real effects of academic research. *American Economic Review, 79 (5)*, 957-970.

Jiwattanakulpaisarn, P., R.B. Noland and D.J. Graham, "Highway infrastructure and private output: evidence from static and dynamic production function models," *Transportmetrica* 7 (2011), 347-367.

Krugman, P. (1991). *Geography and Trade.* Cambridge, MA: MIT Press.

Kulish, M., Richards, A. & Gillitzer, C. 2012. Urban Structure and Housing Prices: Some Evidence from Australian Cities, *Economic Record*, 88, 303-322.

Landcare Research and MAF (2002). Land-use capability map obtained by Motu.

Lucas, R. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22(1)*, 3-42.

Lutchman, L. (2013). *Transport Links and New Zealand Regional Growth: A Long Run Analysis*. Bacelor of Commerce (Honours) Dissertation, University of Auckland.

Lynn, I.H., Manderson, A.K., Page, M.J., Harmsworth, G.R., Eyles, G.O., Douglas, G.B, Mackay, A.D., Newsome, P.J.F. (2009). *Land Use Capability Survey Handbook – a New Zealand handbook for the classification of land* 3rd ed. Hamilton: AgResearch; Lincoln: Landcare Research; Lower Hutt: GNS Science.

Mare D.C., Graham DJ. 2013. "Agglomeration elasticities and firm heterogeneity", *Journal of Urban Economics*, 75, 44-56.

McCann, P. (2013). *Modern Urban and Regional Economics*. Oxford: Oxford University Press.

Mera, K., "II. Regional production functions and social overhead capital: An analysis of the Japanese case," *Regional and Urban Economics* 3 (5// 1973), 157-185.

E S Mills. An Aggregative Model of Resource Allocation in a Metropolitan Area. *American Economic Review*, Papers and Proceedings, 57(2), 197-210, 1967.

R F Muth. *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. Third Series: Studies in Business and Society. University of Chicago Press, Chicago, 1969.

Mollick, A. V., & Mora, M. T. (2012). The impact of higher education on Texas population and employment growth. *The Annals of Regional Science, 48(1)*, 135-149.

Motu Economic and Public Policy Research (2013). *A New Zealand Urban Population Database.* http://www.motu.org.nz/building-capacity/dataset/new_zealand_urban_population_data

Munnell, A.H., "How Does Public Infrastructure Affect Regional Economic Performance?," *New England Economic Review* (1990), 11-32.

National Library of New Zealand (2014a). Census of the Maori Population, 1881 (Papers Relating to the), *Appendix to the Journals of the House of Representatives, 1881 Session I, G-03*. Retrieved from http://atojs.natlib.govt.nz/

National Library of New Zealand (2014b). Railway Commission (Report of.), *Appendix to the Journals of the House of Representatives, 1880 Session I, E-03.* Retrieved from http://atojs.natlib.govt.nz/

National Library of New Zealand (2014c). Railways Statement. (4th September, 1901.) By the Minister for Railways, the Hon. Sir J.G. Ward, K.C.M.G., *Appendix to the Journals of the House of Representatives, 1901 Session I, D-02.* Retrieved from http://atojs.natlib.govt.nz/

New Zealand Ministry of Education. (2014, June 9). *More Young People with NCEA Level 2*. Retrieved from New Zealand Ministry of Education Web site: http://www.minedu.govt.nz/theMinistry/BetterPublicServices/More18YearOldsWithNCEALevel2.asp

New Zealand Railways (1937). *A Record of Progress 1937.*

New Zealand Railways (1957). *Geographical Mileage Table 1957.*

NIWA (2014). *The National Climate Database.* Available at: http://cliflo.niwa.co.nz/

Overman, H., Rice, P., Venables, A. 2010. Economic Linkages Across Space. *Regional Studies* 44(1), 17-33.

Rappaport, Jordan. 2007. “Moving to Nice Weather.” *Regional Science and Urban Economics*, 37 (3): 375–98.

Rappaport, Jordan. 2008. “Consumption Amenities and City Population Density.” *Regional Science and Urban Economics*, 38 (6): 533–52.

Rappaport, Jordan. 2009. “The Increasing Importance of Quality of Life.” *Journal of Economic Geography*, 9 (6): 779–804.

Rappaport, Jordan, and Jeffrey D. Sachs. 2003. “The United States as a Coastal Nation.” *Journal of Economic Growth*, 8 (1): 5–46.

Rodriguez-Pose, A. & Crescenzi, R. (2008). "Research and Development, Spillovers, Innovation Systems, and the Genesis of Regional Growth in Europe," *Regional Studies*, 42(1), 51-67.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy, 98(5)*, 71-102.

Shadbolt, M. (1968). *The Shell Guide to New Zealand*, Christchurch: Whitcombe and Tombs.

Statistics New Zealand (various). *The New Zealand Official Yearbook*, various years. Retrieved from the SNZ Yearbook collection:1893-2012, http://www.stats.govt.nz/browse_for_stats/snapshots-of-nz/digital-yearbook-collection.aspx.

Sterlacchini, A. (2008). R&D, higher education and regional growth: Uneven linkages among European regions. *Research Policy, 37(6-7)*, 1096-1107.

Te Ara Encyclopedia of New Zealand (2014). *Universities before 1900*, Retrieved from http://www.teara.govt.nz/en/tertiary-education/page-1

Trendle, B., & Warren, P. (2004). The role of education in regional income determination - a cross sectional study of small areas in Queensland. *Australasian Journal of Regional Studies, 10(3),* 383.

Wang, H. (2010). *Institutions of higher education and the regional economy: A long-term spatial analysis* (Report ID. 376148). Economics Research International: Hindawi Publishing.

Wu, J.J., Gopinath, M., 2008. What Causes Spatial Variations in Economic Development in the United States? *American Journal of Agricultural Economics* 90(2), 392–408.

Yu, N., M. de Jong, S. Storm and J. Mi, "Spatial spillover effects of transport infrastructure: evidence from Chinese regions," *Journal of Transport Geography* 28 (4// 2013), 56-66.

Appendix

TABLE A1: ***Data Sources and Sample Means***

Variable	Mean	Source
Dependent variable:		
Geometric average annual population growth in decade *t (%)*	1.19	Motu (2013)
Independent variables:		
Average land-use capability	2.89	Landcare Research and MAF (2002)
Average annual sunshine hours	1996	NIWA (2014)
1932 road distance to Auckland (miles, Auckland region only)	87.08	Alexander Turnbull Library (2006)
1932 road distance to Wellington (miles, Wellington region only)	72.50	Alexander Turnbull Library (2006)
1932 road distance to Christchurch (miles, Christchurch region only)	56.40	Alexander Turnbull Library (2006)
1932 road distance to Dunedin (miles, Dunedin region only)	65.43	Alexander Turnbull Library (2006)
1968 road distance to Auckland (miles, Auckland region only)	80.83	Shadbolt (1968: p33)
1968 road distance to Wellington (miles, Wellington region only)	70.38	Shadbolt (1968: p33)
1968 road distance to Christchurch (miles, Christchurch region only)	52.60	Shadbolt (1968: p33)
1968 road distance to Dunedin (miles, Dunedin region only)	64.43	Shadbolt (1968: p33)
Region		
Auckland	0.21	
Wellington	0.14	
Christchurch	0.09	
Greater Auckland	0.09	
Greater Wellington	0.23	
Greater Christchurch	0.11	
Dunedin	0.13	
Percentage Maori 1881	13.69	National Library of New Zealand (2014a)
Percentage Maori 1946	1.38	Motu (2013)
Log of Start of Decade Population	9.07	Motu (2013)
Distance to port in 1932 (miles)	38.57	Alexander Turnbull Library (2006)
University at start of decade *t* (dummy)	0.10	Te Ara Encyclopedia of New Zealand (2014)
Polytechnic at start of decade *t* (dummy)	0.23	Dougherty (1999)
Airport in 1966 (dummy)	0.41	SNZ (various)
Average TLA “life” ranking, 1996-2006		Donovan (2011)
TLA “business” ranking, 1996-2006		Donovan (2011)
Instruments:		
1880 has rail (dummy)	0.63	National Library of New Zealand (2014b)
1880 rail to be delayed (dummy)	0.13	National Library of New Zealand (2014b)
1880 rail to be prioritised (dummy)	0.07	National Library of New Zealand (2014b)
1901 on coach route (dummy)	0.29	National Library of New Zealand (2014c)
1901 population	8570	Motu (2013)
1901 population data missing (dummy)	0.30	Motu (2013)
1901 railroad (dummy)	0.80	National Library of New Zealand (2014c)
1901 steamer port (dummy)	0.18	National Library of New Zealand (2014c)
1906 port (dummy)	0.29	SNZ (various)

TABLE A1 (cont'd): *Data Sources and Sample Means*		
Variable	**Mean**	**Source**
Instruments (continued):		
1906 port tonnage	351395	SNZ (various)
1909 rail distance to Auckland (miles, Auckland region only)	73.26	New Zealand Railways (1937, 1957)
1909 rail distance to Christchurch (miles, Christchurch region only)	51.79	New Zealand Railways (1937, 1957)
1909 rail distance to Dunedin (miles, Dunedin region only)	66.02	New Zealand Railways (1937, 1957)
1909 rail distance to Wellington (miles, Wellington region only)	73.58	New Zealand Railways (1937, 1957)
1920 has a dairy factory (dummy)	0.79	Alexander Turnbull Library (2006)
1920 has a meatworks (dummy)	0.36	Alexander Turnbull Library (2006)
1926 university (dummy)	0.07	SNZ (various)
40-year population lag (in logs)	8.73	Motu (2013)
Average annual rainfall (mm)	1190	NIWA (2014)
Average summer max temperature (degrees Celsius)	21.9	NIWA (2014)
Average winter max temperature (degrees Celsius)	13.0	NIWA (2014)
Straight-line distance to Auckland (km, Auckland region only)	100.80	GPS Visualizer (2014)
Straight-line distance to Christchurch (km, Christchurch region only)	78.01	GPS Visualizer (2014)
Straight-line distance to Dunedin (km, Dunedin region only)	83.21	GPS Visualizer (2014)
Straight-line distance to Wellington (km, Wellington region only)	92.79	GPS Visualizer (2014)
Other data:		
Hospital in 1916 (dummy)		SNZ (various)
Hospital in 1926 (dummy)		SNZ (various)
Hospital admissions 1915		SNZ (various)
Hospital admissions 1924		SNZ (various)
1968 road travel time to Auckland		Shadbolt (1968: p326)
1968 road travel time to Wellington		Shadbolt (1968: p326)
1968 road travel time to Christchurch		Shadbolt (1968: p326)
1968 road travel time to Dunedin		Shadbolt (1968: p326)
Total port tonnage (1916-1976)		SNZ (various)
Total number of vessels to port (1926-1976)		SNZ (various)
Aerodrome in 1936 (dummy)		SNZ (various)
Regular commercial flights (dummy, 1946-1956)		SNZ (various)
Port in 1903 (dummy)		SNZ (various)
Port (dummy, 1916-1976)		SNZ (various)
International flights (dummy, 1946-2006)		SNZ (various)
Permanent & long-term arrivals to NZ in decade *t*		SNZ (various)
Not connected to a main centre by rail in 1909 (dummy)		New Zealand Railways (1937, 1957)

Notes: NIWA (2014) data were not available for 18 towns. In these cases, climate data were approximated with the values of the nearest neighbouring town.

TABLE A2:
Summary Statistics -Average Annual Population Growth by Decade

Year	Mean	SD	Min	Max
1936	1.08	1.10	-1.17	4.67
1946	0.95	1.12	-0.59	4.76
1956	3.04	1.69	-1.98	7.34
1966	2.69	2.45	-1.08	12.69
1976	1.45	1.35	-0.75	4.39
1986	0.36	0.94	-1.41	3.29
1996	0.08	0.94	-2.10	2.88
2006	-0.10	1.06	-2.25	3.31
All years	1.19	1.77	-2.25	12.69

TABLE A3:
Region Classifications and Road Distance to Each Main Centre

AUCKLAND

	Distance to Auckland (km)
Auckland	*0*
Pukekohe	48
Huntly	100
Hamilton	134
Paeroa	137
Morrinsville	145
Cambridge	156
Waihi	159
Te Aroha	161
Te Awamutu	164
Whangarei	174
Dargaville	183

GREATER AUCKLAND

	Distance to Auckland (km)
Te Kuiti	214
Tauranga	220
Rotorua	245
Taumaranui	299
Whakatane	320

WELLINGTON

	Distance to Wellington (km)
Wellington	*0*
Carterton	90
Levin	100
Masterton	103
Foxton	119
Palmerston North	150
Feilding	172
Marton	172

GREATER WELLINGTON

	Distance to Wellington (km)
Wanganui	203
Dannevirke	212
Waipukurau	269
Hawera	298
Ohakune	299
Hastings	317
Eltham	319
Stratford	330
Napier	341
New Plymouth	370
Waitara	370
Wairoa	468
Gisborne	575

CHRISTCHURCH

	Distance to Christchurch (km)
Christchurch	*0*
Rangiora	32
Ashburton	87
Temuka	145
Timaru	159

GREATER CHRISTCHURCH

	Distance to Christchurch (km)
Waimate	209
Greymouth	254
Hokitika	259
Blenheim	323
Westport	346
Nelson	441

DUNEDIN

	Distance to Dunedin (km)
Dunedin	*0*
Milton	55
Balclutha	80
Kaitangata	90
Oamaru	117
Gore	159
Invercargill	224

TABLE A4:	**Dependent variable: Decade average annual % population growth**					
Pooled OLS estimates with quadratic in TLA amenity rankings	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	1926-2006		**1926-1966**		**1966-2006**	
Average land-use capability	0.285***	0.275***	0.362**	0.355**	0.210***	0.194***
	(0.101)	(0.102)	(0.156)	(0.164)	(0.073)	(0.071)
Average annual sunshine hours (00s)	0.260***	0.265***	0.280***	0.281***	0.238***	0.247***
	(0.042)	(0.044)	(0.060)	(0.062)	(0.038)	(0.038)
Initial road miles to Auckland (00s)	-0.287	-0.271	0.102	0.041	-0.862	-0.772
	(0.605)	(0.607)	(0.783)	(0.783)	(0.696)	(0.715)
Initial road miles to Wellington (00s)	0.329	0.390	0.096	0.102	0.667	0.793*
	(0.495)	(0.452)	(0.610)	(0.641)	(0.547)	(0.458)
Initial road miles to Christchurch (00s)	-0.128	-0.106	0.016	-0.024	-0.287	-0.194
	(0.407)	(0.413)	(0.361)	(0.354)	(0.864)	(0.891)
Initial road miles to Dunedin (00s)	0.266	0.306	0.582	0.529	-0.220	-0.087
	(0.247)	(0.244)	(0.395)	(0.386)	(0.283)	(0.291)
Region (omitted category: Auckland)						
Wellington	-0.787	-0.871	-0.030	-0.119	-1.701**	-1.802***
	(0.625)	(0.597)	(0.837)	(0.900)	(0.686)	(0.639)
Christchurch	-0.558	-0.640	-0.457	-0.547	-0.773	-0.869
	(0.553)	(0.545)	(0.622)	(0.686)	(0.835)	(0.833)
Greater Auckland	0.902	0.836	2.026**	2.121***	-0.343	-0.579
	(0.584)	(0.597)	(0.760)	(0.782)	(0.663)	(0.717)
Greater Wellington	-1.045*	-0.993*	-0.737	-0.813	-1.459**	-1.284*
	(0.541)	(0.550)	(0.640)	(0.639)	(0.665)	(0.714)
Greater Christchurch	-0.665	-0.741	-0.064	-0.160	-1.250*	-1.327**
	(0.574)	(0.555)	(0.722)	(0.765)	(0.689)	(0.659)
Dunedin	-0.806	-0.880	-0.362	-0.429	-1.196*	-1.296*
	(0.563)	(0.545)	(0.730)	(0.758)	(0.677)	(0.656)
Percentage Maori 1881	-0.003		0.006		-0.011***	
	(0.003)		(0.006)		(0.003)	
Percentage Maori 1946		-0.070		0.027		-0.170**
		(0.077)		(0.145)		(0.078)
Log of Start of Decade Population	0.109*	0.112*	0.132	0.126	0.148**	0.158**
	(0.061)	(0.062)	(0.095)	(0.094)	(0.058)	(0.060)
Average amenity ranking	0.007	0.007	0.024	0.028	-0.001	-0.005
	(0.025)	(0.025)	(0.037)	(0.037)	(0.019)	(0.020)
Average amenity ranking squared	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	448	448	224	224	224	224
R-squared	0.598	0.598	0.529	0.528	0.631	0.624

All regressions include an intercept and time fixed effects. Standard errors clustered on town in parentheses.

* p<.1, ** p<.05, *** p<.01

TABLE A5: *Pooled OLS estimates with quadratic in TLA business rankings*	**Dependent variable: Decade average annual % population growth**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	1926-2006		**1926-1966**		**1966-2006**	
Average land-use capability	0.211*	0.207*	0.287	0.288	0.182**	0.166*
	(0.116)	(0.120)	(0.177)	(0.186)	(0.085)	(0.088)
Average annual sunshine hours (00s)	0.250***	0.254***	0.272***	0.267***	0.228***	0.238***
	(0.041)	(0.043)	(0.059)	(0.062)	(0.037)	(0.037)
Initial road miles to Auckland (00s)	0.024	0.068	0.535	0.432	-0.935	-0.750
	(0.717)	(0.725)	(1.013)	(1.032)	(0.697)	(0.743)
Initial road miles to Wellington (00s)	0.611	0.649	0.368	0.313	0.750	0.890
	(0.689)	(0.686)	(0.952)	(0.992)	(0.575)	(0.545)
Initial road miles to Christchurch (00s)	0.042	0.060	0.235	0.195	-0.252	-0.174
	(0.498)	(0.509)	(0.555)	(0.556)	(0.833)	(0.896)
Initial road miles to Dunedin (00s)	0.103	0.113	0.578*	0.568*	-0.365**	-0.332*
	(0.152)	(0.151)	(0.290)	(0.290)	(0.172)	(0.176)
Region (omitted category: Auckland)						
Wellington	-0.468	-0.462	0.405	0.328	-1.687***	-1.643***
	(0.561)	(0.572)	(0.839)	(0.893)	(0.547)	(0.549)
Christchurch	-0.158	-0.136	0.066	-0.048	-0.788	-0.679
	(0.630)	(0.631)	(0.800)	(0.820)	(0.767)	(0.823)
Greater Auckland	1.231*	1.197*	2.452***	2.527***	-0.354	-0.516
	(0.636)	(0.634)	(0.885)	(0.908)	(0.651)	(0.713)
Greater Wellington	-0.596	-0.536	-0.187	-0.313	-1.430**	-1.197
	(0.673)	(0.686)	(0.939)	(0.971)	(0.663)	(0.740)
Greater Christchurch	-0.222	-0.191	0.350	0.206	-1.167*	-1.015
	(0.614)	(0.614)	(0.894)	(0.917)	(0.644)	(0.670)
Dunedin	-0.070	-0.038	0.254	0.116	-0.889	-0.742
	(0.786)	(0.777)	(1.159)	(1.164)	(0.694)	(0.709)
Percentage Maori 1881	-0.003		0.007		-0.011***	
	(0.003)		(0.006)		(0.004)	
Percentage Maori 1946		-0.034		0.049		-0.126**
		(0.064)		(0.130)		(0.061)
Log of Start of Decade Population	0.076	0.082	0.113	0.098	0.103	0.130
	(0.076)	(0.076)	(0.107)	(0.105)	(0.074)	(0.078)
Average business ranking	-0.020	-0.022	-0.029	-0.025	0.007	0.000
	(0.042)	(0.042)	(0.057)	(0.058)	(0.031)	(0.032)
Average business ranking squared	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
N	448	448	224	224	224	224
R-squared	0.600	0.600	0.529	0.527	0.640	0.629

All regressions include an intercept and time fixed effects. Standard errors clustered on town in parentheses.

* p<.1, ** p<.05, *** p<.01