THE
ECONOMIC
JOURNAL

The Economic Journal, 1–34 DOI: 10.1093/ej/ueaa088
Advance Access Publication Date: 11 July 2020

HOUSE PRICES AND GROWTH WITH FIXED LAND SUPPLY*

David Miles and James Sefton

We analyse housing costs and patterns of residential development over the long term in a dynamic general equilibrium. We show that in a growing economy the speed of travel improvements is crucial to the evolution of land and house prices. We derive a condition for the rate of change in transport efficiency that generates flat land and house prices on a balanced growth path. We present evidence that this condition was satisfied in many countries between the mid-nineteenth century and the mid-twentieth century, but since then passenger transport improvements have slowed down with major implications for how house prices evolve.

Across developed economies on average house prices hardly changed relative to the price of consumer goods in the period between the mid-to-late nineteenth century and the mid-twentieth century. The left hand panel of Figure 1 shows an index of average real house prices for a group of 16 developed countries between 1870 and 2016 using an updated version of the data from Knoll *et al.* (2017). The real house price index from these developed countries, adjusted by national consumer prices and averaged across all countries is—net—flat between 1870 and 1950. But between 1950 and 2016 real house prices then quadrupled. The rises were particularly marked in the period since 1990. This pattern of almost 100 years with average real house prices overall unchanged, then rising steadily before a more recent period of very rapid price rises is the mirror image of what has happened to commuting speeds. As documented in the Appendix to this paper, between the middle of the nineteenth century and up to the World War II the average speed of commuting (and commuting distances) more than tripled in the most advanced economies as rail travel spread widely and supplanted horses and walking as the main means of getting to work. In the period between the World War II and 1970 travel speeds rose further but at a less rapid rate as automobiles took over from trains as the main means of transport in most developed countries. But the speed advantage of road over rail for commuter trips was far less than the advantage of rail over horses and of walking. And since 1980 speed improvements have largely stopped—they may even be in decline in many developed world cities as congestion has increased. In the USA, average commuter travel speeds in 1977 were 35 mph but by 2017 had fallen to around 27 mph (US Department of Transport); in England between 1988 and 2014 there was little change in the

* Corresponding author: David Miles, Imperial College Business School, Imperial College London, South Kensington Campus, London SW7 2AZ, UK. Email: d.miles@imperial.ac.uk

This paper was received on 21 March 2019 and accepted on 3 July 2020. The Editor was Francesco Lippi.

The data and codes for this paper are available on the Journal website. They were checked for their ability to reproduce the results presented in the paper.

An editor and three referees provided many detailed, helpful comments on an earlier draft of this paper. The authors would like to thank seminar participants at the Bank of England, Cambridge University, Canada Housing and Mortgage Corporation, Edinburgh University, Imperial College London, Oxford University, Warwick University, the UK Department of Communities and Local Government and HM Treasury for helpful comments. The authors would also like to particularly thank both Alex Blei and Shlomo Angel from the NYU Urban Expansion Program at the Marron Institute of Urban Management and the Stern School of Business of New York University for providing data on urban expansion over the last two centuries. A previous version of this paper circulated with the title 'Houses across Time and across Place'. James Sefton is supported by the Economic and Social Research Council grant number ES/P000142/1.

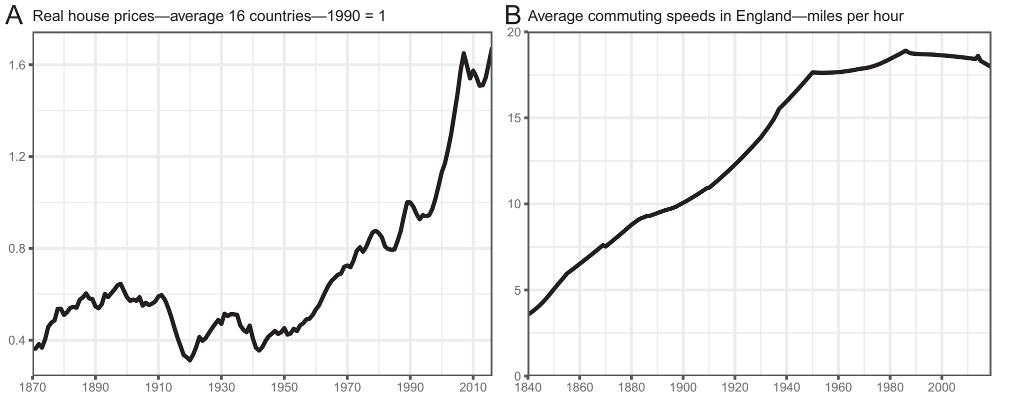

Fig. 1. *House Prices and Commuting Speeds.*
Source. The house price data are from the Jorda-Schularick-Taylor Macrohistory Database (available at http://www.macrohistory.net/data/). Details of the data and approach used to construct estimates of commuting speeds in England are in Appendix A.

average distance travelled by workers to their place of work, but a near 30% rise in journey times (Commuting Trends in England 1988–2015, UK Department of Transport).

The right-hand panel of Figure 1 shows an estimate of passenger transport speeds over the past 180 years in England—the country that for most of the earlier part of this period was at the technological frontier as regards improvements in passenger travel. Between 1840 and 1930 we estimate that travel speeds in England rose by around 1.5% a year; this then slowed to around 0.6% a year up to 1980 since when travel speeds have slightly fallen. In England the years of peak improvements in speed came a little earlier than in most other developed countries though the improvements were ultimately comparable in other countries.

In this paper we explore the link between the phenomena—illustrated in Figure 1—of periods of fast rises in travel speeds and flat house prices followed by sharply slowing travel improvements and sustained and large rises in real house prices. We show that with a plausible model of urban expansion in which there are advantages to workers and consumers of being close to the urban centre, the change in house prices over time is sensitive to the change in the speed at which people can get close to the heart of the city. We use a framework that combines features of a Ramsey two-sector growth model with a model of the changing geography of residential development that tracks the change in location of the population over time. We derive a condition under which in a growing economy the speed at which transport technology (commuting speed) improves relative to the growth in aggregate incomes is such as to generate flat house prices. The condition is simple and intuitive and links the speed of travel improvements to one-half the growth in GDP—a link that arises because simple geometry means that the area that it is feasible for commuting into central places grows at twice the rate of increase in travel speeds. We present evidence that in the USA this condition seems close to holding in the period when real house prices appear to have been (on average) flat—that is the period from the late nineteenth to the mid-twentieth century.

The model we develop both helps explain the past and has implications for how house prices might evolve over the next several decades. For the past, we are able to explain why real house prices were flat across many countries for a long period but then rose sharply, but at rates that differed between countries. We find the factor that was important for explaining house prices rises in recent decades is the interaction between a sharply slowing rate of improvement in transport speeds (common across developed countries) and the average availability of land to develop around cities (which differed between countries). For the future, our model suggests that in the absence of a return to rapid rises in travel speeds, and unless price elasticities of demand for housing are at the high end of a plausible range, then in those countries with more limited undeveloped land house prices will tend to rise faster than the prices of other goods and may consistently outpace income growth.

One of the contributions of this paper is to develop a model that is tractable and which can be used to explore how alternative technological developments are likely to impact housing market outcomes and the ways in which their effects are sensitive to consumer preferences and to house building techniques. We show how a slowing in the pace of transport improvements is likely to generate house price rises that over time can be very substantial. The model helps explain the main differences in the house price history of the USA—a land-rich country—relative to the UK, a far more densely populated country, over the period when commuting speed improvements slowed in both countries and then stopped. Shiller's data suggest US real house prices were roughly flat between the end of the nineteenth century and 1950 but have risen by about 70% since then (less than the rise in per capita real income); in the UK real house prices are well over four times as high in 2018 than just after the World War II and have risen greatly relative to incomes.

As well as using the model to help understand the similarities and differences between changes in real house prices over the past 150 years in the USA and in the UK, we analyse possible paths of home prices over the next 50 years and how they are affected by two key elasticities reflecting substitutability between land and structures in creating housing and between housing and other goods in generating utility.

Related literature

There is a large literature on the determinants of house prices, both across time and across regions within countries. Nearly all of the literature focuses either on the time dimension of average prices or on the cross section of prices by location. There is another large literature on the allocation of people and economic activity across space. (For a good review of the literature on many aspects of housing economics, see the survey papers by Redding, 2013 and Piazzesi and Schneider, 2016; on spatial economics an important work is Fujita *et al.*, 2001). Much of the literature focuses on the variability in land and house prices, and in the density of development and populations, across regions at a point in time. The pioneering works are Alonso (1960; 1964), Mills (1967; 1972) and Muth (1969); a resurgence in the literature was triggered by Krugman (1991) and Lucas (2001). We introduce variability in housing, location and land values in a tractable way that captures the essence of these ideas, which is that there are desirable locations and that being further from them creates costs. Land and house prices adjust to reflect that. We map out how this adjustment evolves over time and in so doing model the overall macroeconomy and its path over long periods. The focus of much of the literature on national house prices is less on the very long-term drivers of housing markets and more on business cycle variability in values. Much of the literature on regional differences in housing conditions does not focus on the macroeconomic backdrop so takes aggregate incomes, interest rates and population as given (and often constant). A substantial literature looks at how housing fits into household decisions on portfolio allocation, borrowing and saving (see, for example, Campbell and Cocco, 2003;

2007). In much of this literature the changing way in which housing is supplied is not the focus of attention; supply is often assumed fixed or at least exogenous. In contrast our focus is on the long run and on how the amount and location of housing evolves.

One paper in the spirit of our own is Deaton and Laroque (2001); see also Kiyotaki *et al.* (2011); Grossmann and Steger (2016); Favilukis *et al.* (2017). Those papers, like this one, embed the housing sector within a model of the overall economy that endogenises growth, saving and asset prices. But they do not focus on the spatial dimension of housing, which we show is important when thinking about long-run dynamics. Any long-run analysis has to model the changing supply of housing taking into account the fixity of land mass and the way in which endogenous shifts in the cost of land relative to structures changes the way in which houses are constructed. Land is obviously not homogeneous and the impact of the most important way in which it differs (that is by location) varies over time as technological change means that distance may have a varying effect on value. One obvious way in which this happens is if transport costs change. That is central to our result on balanced growth. We show there is a balanced growth path (BGP) with constant land and house prices for general specifications of the elasticity of substitution between land and structures even though land is in fixed supply. This is because technical progress in transport effectively increases the amount of land where residential development is feasible. Other models with fixed land can generate steady paths but with changing relative prices and often with specific assumptions on elasticities to stop spending shares going to zero or one (e.g., Kiyotaki *et al.*, 2011). Our result does not depend on a restriction of unit elasticities, a restriction that is not consistent with a great deal of evidence.

A number of papers have developed urban growth models that admit BGPs. These papers, though, are primarily focused on explaining the number of cities rather than describing the urban footprint. In both Black and Henderson (1999) and Rossi-Hansberg and Wright (2007) the rising costs of commuting in a monocentric city are offset by scale or knowledge spillovers due to greater agglomeration. In Davis *et al.* (2014) land prices rise along their BGP; indeed it is these rising land prices that lead to increases in the production density that result in productivity increases from agglomeration. Their paper treats the period since 1970 as one of balanced growth, whereas we model the period before 1970—when land prices were roughly constant—as a BGP, and treat the period since 1970 when land prices have risen as being off the BGP. Rossi-Hansberg and Wright (2007) also analyse the evolution of cities where rising agglomeration benefits from larger city size are traded off against rising transport costs. In their model new cities can be formed and in that sense there is no constraint on land availability whereas in our model we allow for land availability to be more of a constraint so that balanced growth requires that improvements in commuting technology come at a rate that depends on growth in population and in per capita real incomes.

Heblich *et al.* (2018) makes the evolution of passenger transport central to the rise of the first mega city in the nineteenth century—London. Like this paper it makes speed of commuting a driver of the path of house values.

1. The Model

1.1. *The Physical Environment*

We assume a circular economy in which people live at locations of varying distance $l \in [0, l_{\max}]$ from the centre ($l = 0$) potentially up to the periphery of the economy (at $l_{\max}$). The physical area of the economy is $\pi l_{\max}^2$ which is the aggregate quantity of land. The central business district (CBD) is located at $l = 0$. There are advantages to living close to the centre. One interpretation is that this is where the jobs are; another is that this is where the best choice of goods is available. Though we recognise that in actual economies there may be many more than one central business location, we adopt the assumption of a monocentric economy for analytical tractability. Key results do not depend crucially on there only being one urban centre. We expand on this below.

The monocentric assumption allows us to describe the distance to work, or to the widest choice of goods, from anywhere in the economy by the distance l to the centre, implying that the urban area where all housing is distance l from the centre is an annulus. This means it is easy to integrate the populated area of the economy. It also implies that the rate of expansion in the urban area will be linear in its radius; this will prove useful for demonstrating the existence of a BGP. A more involved geography with many urban centres would make these expressions more complicated once cities begin to overlap (we discuss this extension further in Subsection 1.8). However, the core ideas in this paper will all go through as long as all jobs pay the same wage so that all else equal people want to locate at their closest CBD. One could even allow for emergence of new CBDs, though this would most likely lead to discontinuities.

We assume a simple cost of distance function levied on consumption. Thus, we follow a large literature (e.g., Krugman, 1980) in using Samuelson's 'Iceberg' model of transport costs; that a fraction of any good shipped simply 'melts away' in transit. Formally, at distance l from the CBD, $1 + \lambda_t l$ of consumption good must be purchased to consume 1 unit of the good. We think of $\lambda_t l$ as the tax on location with λ_t as the impact at time t of distance on that tax rate.[1] There is no reason to think that $\lambda_t l$ is constant over time—it reflects technology (most obliviously travel technology) which has improved dramatically over the last couple of centuries. We come back to this shortly.

1.2. *The Households*

The agents in the model are households each of which is a member of an infinitely lived dynasty. There is a continuum of dynasties on the unit interval. Though the number of dynasties remains constant over time, the number of people in the current household of each dynasty grows at rate m which is therefore also the rate of population growth. If we normalise total population to be 1 at time $t = 0$, then at time t the population, $n(t)$, is equal to e^{mt}. Labour is supplied inelastically by each household in proportion to dynastic size at each period, and so the labour force, L_t, grows at rate m too.

Each dynasty derives utility from the consumption of goods, denoted C, and of housing services, S. Preferences over these goods at a given time t is described by a constant elasticity of substitution (CES) utility function

$$Q_{it} = \left[aC_{it}^{1-1/\rho} + (1-a)S_{it}^{1-1/\rho} \right]^{1/(1-1/\rho)},$$

[1] There are many aspects of the cost of location and several interpretations of $\lambda_t l$. The most obvious is travel costs—you need to spend time and money on getting nearer to the centre where you may work and where you can most easily buy goods and consume them. This idea goes back at least to von Thunen *et al.* (1966). Many goods need to be brought to location l at greater cost than being brought to places nearer the centre and that to go to the centre and buy them cost you time and travel expenses; this is in the spirit of Samuelson's iceberg costs of moving things and the net effect is that the costs of such goods is raised by $\lambda_t l$. It is also consistent with Krugman's model of commuting costs, where all dynasties have a fixed supply of labour but lose a proportion of this supply in commuting to the CBD for work.

where $\rho \in (0, 1)$ is the elasticity of substitution between housing and consumption goods and a is a share parameter. We refer to the quantity Q as the composite consumption good or simply the composite good. The indices $i \in [0, 1]$ and $t \in [0, \infty)$ index the quantity to dynasty i at time t. Dynastic welfare is the discounted power function of the composite good

$$\int_0^\infty \frac{1}{1-\gamma} Q_{it}^{(1-\gamma)} e^{-\theta t} dt, \tag{1}$$

where γ reflects the degree of intertemporal substitutability and θ is the discount factor. We use the dynasty[2] as our decision-making unit throughout.

Dynasties can choose where to live at each point in time. There are no costs to moving from one location to another so that dynasties, given a chosen expenditure, pick the location that maximises the consumption of the composite good at each point in time. As dynasties are homogeneous,[3] if there was one location preferred by a dynasty then all dynasties would prefer this same location. Therefore a necessary condition for an equilibrium in the housing market is that the maximum utility that can be derived from each location is the same. In this housing market, the centripetal force attracting dynasties towards the CBD is the lower transport costs. The offsetting centrifugal force is that housing costs (that is rental rates) become more expensive closer to the CBD. As there are no other forces affecting the choice of location—for example some preference for a less densely populated environment or alternatively a preference for being close to other people—these two forces need to be exactly offsetting at each location in equilibrium. We shall now flesh out this argument.

The consumption good is the numeraire, the real interest rate is denoted r_t and the rental price of housing services at location l at time t is p_{lt}^S (also referred to as the user cost of housing). Define the composite price at location l as p_{lt} where $p_{lt} = \left(a^\rho (1+\lambda_t l)^{1-\rho} + \left(p_{lt}^S\right)^{1-\rho} (1-a)^\rho\right)^{1/(1-\rho)}$. We can then write the optimal consumption allocation of dynasty i choosing location l at time t in terms of the composite good and composite price as

$$C_{i(l),t} = \left(\frac{a p_{lt}}{(1+\lambda_t l)}\right)^\rho Q_{it} \text{ and } S_{i(l),t} = \left(\frac{(1-a) p_{lt}}{p_{lt}^S}\right)^\rho Q_{it}, \tag{2}$$

where the notation $C_{i(l),t}$ and $S_{i(l),t}$ explicitly recognises the location of dynasty i. It follows that the total expenditure of this dynasty can then be expressed as the consumption on the composite good times its price

$$(1+\lambda_t l) C_{i(l),t} + p_{lt}^S S_{i(l),t} = p_{lt} Q_{it}. \tag{3}$$

As dynasties are identical and there are no costs to moving, then in equilibrium all dynasties must have the same utility at all t. All dynasties, therefore, have the same composite consumption, Q_{it}; though of course the allocation of each dynasty is different and depends on their location. We

[2] We could equally have done the analysis in per capita terms. Assume that the flow of dynastic utility at time t is the sum of utilities of identical dynastic members then alive—whose number is proportional to $n(t)$. Because the utility function is constant returns to scale (CRS) and population growth is constant the welfare function in (1) can be rewritten in per capita terms but with an adjusted discount rate $\widetilde{\theta} = \theta + \gamma m$. Thus the dynastic welfare function (1) is equivalent to a welfare function that is the sum over members of the dynasty of their individual utilities, but with a shifted discount factor.

[3] This assumption could be weakened. As long as all dynasties have the same CES preferences over the consumption goods, they will choose to consume the two goods—consumption and housing—in the same proportions at each location. So even if there was heterogeneity in wealth holdings across dynasties, the main results in this paper would go through.

denote this level of consumption as $Q_{it} = Q_t$ (which is also equal to total aggregate consumption as dynasties have unit mass). It also implies that the cost of the composite good must be the same at every location;[4] thus $p_{lt} = p_{0t}$ for all l. For this condition to hold, the rental price of housing services at each location l must satisfy

$$p_{lt}^{S} = \left(\left(p_{0t}^{S}\right)^{1-\rho} - \left(\frac{a}{1-a}\right)^{\rho} \left((1+\lambda_t l)^{1-\rho} - 1\right) \right)^{1/(1-\rho)}, \tag{4}$$

where p_{0t}^{S} is the rental price at the CBD. By expressing the cost of travelling as a tax on consumption, we are able to derive this explicit form for what is known in the spatial geography literature as the bid-price function. In the standard statement of the Alonso–Mills–Muth urban model, the cost of travelling is treated as an extra expenditure in addition to the expenditure on consumption and housing. Within this set-up, it is not possible to derive an explicit expression for the bid-price function in all but a few special cases, see Henderson (2014).

Each dynasty is endowed with the same initial wealth, W_0, which will be held in both non-housing physical capital, K, residential structures, B, as well as land, LW. As there is no uncertainty, all assets will deliver the same return, r_t, and so dynasties have no preference over the different asset portfolios. We therefore assume that each dynasty initially holds the same portfolio. The dynastic intertemporal budget constraint at time $t = 0$ can then be written as

$$W_0 + \int_0^{\infty} e^{-\int_0^t r_\tau d\tau} w_t e^{mt} dt = \int_0^{\infty} e^{-\int_0^t r_\tau d\tau} p_{0t} Q_t dt, \tag{5}$$

where the second term is the dynastic human capital: the total present discounted value of future labour income from the supply of one unit of labour by each member of the dynasty at wage w_t. Each dynasty chooses the path for the consumption of the composite good so as to maximise their welfare in (1) subject to their budget constraint (5) and a path for prices r_t, w_t and p_{0t}.

1.3. *Production*

The production side of the economy consists of two sectors; a goods production sector and a housing production sector. The goods production sector uses Cobb–Douglas technology, F, to manufacture the single good. This good can be consumed, C, or invested in productive capital, I_t^K, or in residential buildings, I_t^B. We assume a constant rate of labour augmenting technical progress, g. Thus production in the goods sector is

$$C_t + I_t^K + I_t^B = F\left(K_t, L_t e^{gt}\right) = A K_t^{\alpha} \left(L_t e^{gt}\right)^{1-\alpha}, \tag{6}$$

where α is the capital share of output. All the variables in equation (6) are aggregates; we use the notation that aggregate quantities are indexed by t only. We assume that improvements in building structures proceeds at the same rate as general productivity improvements in goods production; this is a strong assumption but it is not obvious that we have got better at building high buildings more or less rapidly than we have got better at making cars or telephones. The stock of capital, K, and residential buildings, B, evolve over time as

[4] We could derive this condition more formally. As there are no costs to moving, dynasties choose their location at each point in time so as to maximise their utility given their chosen expenditure; (2) and (3) imply that the maximisation problem of dynasty i at time t can be written as choosing l to maximise Q_{it} subject to $p_{lt} Q_{it} \leq E_{it}$ (where E_{it} is the chosen level of expenditure). As the centripetal and centrifugal forces act only through the price, p_{lt}, a necessary condition for an internal maximum is $\partial p_{lt} / \partial l = 0$.

$$\dot{K}_t = I_t^K - \delta^K K_t, \tag{7}$$

$$\dot{B}_t = I_t^B - \delta^B B_t, \tag{8}$$

where δ^K and δ^B are the respective constant depreciation rates of productive capital and residential buildings.

Housing at location l at time t is provided by combining structures, B_{lt}, and land, R_{lt}. The same CES technology is used at all locations, though the mix of buildings and land varies by location. The production of housings services[5] S_{lt} at location l is

$$S_{lt} = H(B_{lt}, R_{lt}) = A_s \left[b B_{lt}^{1-1/\varepsilon} + (1-b) R_{lt}^{1-1/\varepsilon} \right]^{1/(1-1/\varepsilon)},$$

where ε is the elasticity of substitution between land and structure and b is a share parameter. A_s is a constant of proportionality between the flow of services and the stock of housing. As the equilibrium solution is invariant to the product of rental prices, p_{lt}^S, and this flow constant—that is, for example, a doubling of the flow constant leads to a halving of rental prices in equilibrium—we set this constant to 1 in the simulations.

1.4. *Allocation in the Housing Market*

In equilibrium, residential buildings earn a real rate of return, r_t, whatever their location. This condition sets the mix of residential structures to land at each location l.[6] The real return to structures at location l is their marginal product minus depreciation

$$r_t = \left(p_{lt}^S \frac{\partial H(B_{lt}, R_{lt})}{\partial B_{lt}} - \delta_B \right) = p_{lt}^S b A_s \left(b + (1-b) \left(\frac{R_{lt}}{B_{lt}} \right)^{1-1/\varepsilon} \right)^{1/\varepsilon/(1-1/\varepsilon)} - \delta_B. \tag{9}$$

This condition (9) implies that the stock of residential structures and the associated flow of housing services per unit area of land at location l is

$$B_{lt} = \left(\frac{1}{(1-b)} \left(\frac{p_{lt}^S b A_S}{r_t + \delta_B} \right)^{1-\varepsilon} - \frac{b}{(1-b)} \right)^{\varepsilon/(1-\varepsilon)} R_{lt}, \tag{10}$$

$$S_{lt} = A_S \left(\frac{1}{(1-b)} - \frac{b}{(1-b)} \left(\frac{r + \delta_B}{p_{lt}^S b A_S} \right)^{1-\varepsilon} \right)^{\varepsilon/(1-\varepsilon)} R_{lt}. \tag{11}$$

As rental rates, p_{lt}^S, fall away from the centre, equation (10) implies that the ratio of land to structures rises.

The edge of the urban sprawl will be defined by *either* the condition that the marginal product of structures must be greater than or equal to the interest rate, r_t, as the ratio of structures to land tends towards 0 *or* that rental prices, p_{lt}^S, must be greater than or equal to 0. If $\varepsilon < 1$ (and there is a great deal of empirical evidence to suggest it is, Muth, 1971; Ahlfeldt and McMillen,

[5] By S_{lt} we refer to the supply of housing services derived from the buildings B_{lt} and land R_{lt} at location l. In contrast S_{it} refers to the use of housing services by dynasty i. We shall relate the two shortly in order to calculate the amount of land occupied by dynasty i.

[6] In some of the simulations discussed later, we place restrictions on the maximum buildings–land ratios (so as to model planning restrictions). In these cases the following optimal allocation condition is not satisfied. Instead the building–land ratios are given, buildings receive rB_{lt} of the rental income with land receiving the residual.

2014) then the first of these constraints is tighter whereas for $\varepsilon \geq 1$ only the latter bites.[7] Hence for $\varepsilon < 1$ the edge of the urban extent of the economy at time t, denoted by $l_{t,\text{Edge}}$, is

$$l_{t,\text{Edge}} = \min\left(l_{\max}, \frac{1}{\lambda_t}\left(\left(1+\left(p_{0t}^{1-\rho}-\left(\frac{(r_t+\delta_B)}{A_s b^{1/(1-1/\varepsilon)}}\right)^{1-\rho}\right)\left(\frac{(1-a)}{a}\right)^{\rho}\right)^{1/(1-\rho)}-1\right)\right), \tag{12}$$

whereas for $\varepsilon \geq 1$ the edge of the urban extent is the slightly simpler expression

$$l_{t,\text{Edge}} = \min\left(l_{\max}, \frac{1}{\lambda_t}\left(\left(1+p_{0t}^{1-\rho}\left(\frac{(1-a)}{a}\right)^{\rho}\right)^{1/(1-\rho)}-1\right)\right). \tag{13}$$

To complete the description of the housing sector, we consider the price at time t of land at a distance l from the centre, denoted p_{lt}^R. At all locations the return to land must be equal to the real interest rate, r_t. This return to land is the sum of its marginal product, $p_{l\tau}^S \frac{\partial H(B_{l\tau}, R_{l\tau})}{\partial R_{l\tau}}$, plus any capital gains implying

$$r_t p_{lt}^R = p_{l\tau}\frac{\partial H\left(B_{l\tau}, R_{l\tau}\right)}{\partial R_{l\tau}} + \dot{p}_{lt}^R. \tag{14}$$

Integrating this relationship forward subject to the standard transversality condition (that the growth in land prices is less than the interest rate in the long run) gives the price of land as the discounted value of all its future land rents, that is

$$p_{lt}^R = \int_t^{\infty} e^{-\int_t^{\tau} r_\upsilon d\upsilon} p_{l\tau}^S \frac{\partial H\left(B_{l\tau}, R_{l\tau}\right)}{\partial R_{l\tau}} d\tau. \tag{15}$$

We can also derive an alternative expression for rental prices (often referred to as the user cost of housing) as the value weighted average of the gross return to structures plus land rental rates. Given the production of housing services is constant returns to scale the value of the output of housing services is equal to the sum of the gross marginal products times the input good. If we substitute out for the marginal products using equations (9) and (14) and rearrange, then

$$p_{lt}^S = (r_t+\delta_B)\left(\frac{B_{lt}}{S_{lt}}\right) + \left(r_t - \frac{\dot{p}_{lt}^R}{p_{lt}^R}\right)\left(\frac{p_{lt}^R R_{lt}}{S_{lt}}\right).$$

Thus the return, r_t, on a 'house', whose value is $B_{lt} + p_{lt}^R R_{lt}$ at location l, is equal to rents, $p_{lt}^S S_{lt}$, minus depreciation on the buildings plus capital appreciation on the land.

1.5. *Aggregation over Space*

To evaluate aggregate quantities in the housing sector, it is easier integrate over location rather than over dynasties. To do this we need an expression for the dynastic density at location l, that is the mass of dynasties living per unit area at location l. To derive such an expression, note that equation (2) gives the consumption of housing services for a dynasty living at location l and equation (11) gives the supply of housing services per unit area, (S_{lt}/R_{lt}), at location l. The ratio

[7] This follows as the former constraint is equivalent to requiring that $p_{lt}^S \geq \left(\frac{(r_t+\delta_B)}{A_s b^{1/(1-1/\varepsilon)}}\right)$ when $\varepsilon < 1$ whereas the latter is the simpler condition that $p_{lt}^S \geq 0$.

of the two, therefore, gives the land area occupied by a dynasty at location l which we shall denote as $R_{i(l),t} = S_{i(l),t}/(S_{lt}/R_{lt})$. The dynastic density is then $1/R_{i(l),t}$ and the total mass of dynasties living in an annulus of width dl a distance l from the centre is $(2\pi l/R_{i(l),t})dl$.

The aggregate consumption of housing services, S_t, is the integral over the surface of the economy of the demand for housing services of a dynasty at location l times the number of dynasties living in the annulus of radius l;

$$S_t = \int_0^{l_{t,\text{Edge}}} S_{i(l),t}\left(\frac{2\pi l}{R_{i(l),t}}\right) dl = \int_0^{l_{t,\text{Edge}}} S_{lt}\left(\frac{2\pi l}{R_{lt}}\right) dl, \tag{16}$$

where the second expression follows directly from the expression for the dynastic density and is equal to the integral of the amount of housing services per unit area over all locations.[8] The clearing condition in the housing market is that all dynasties must be housed

$$1 = \int_0^{l_{t,\text{Edge}}} \left(\frac{2\pi l}{R_{i(l),t}}\right) dl.$$

It is the rental price of the housing, driven by the rental prices at the CBD p_{0t}^S, which adjusts to clear this market. We state this formally in the next section.

Similarly, we can write an expression for the aggregate demand for the consumption good (which must include the distance tax or consumption 'melt') as the integral of the consumption good used over all dynasties

$$C_t = \int_0^{l_{t,\text{Edge}}} (1+\lambda_t l) C_{i(l),t}\left(\frac{2\pi l}{R_{i(l),t}}\right) dl, \tag{17}$$

and the aggregate demand for residential structures as the integral

$$B_t = \int_0^{l_{t,\text{Edge}}} B_{lt}\left(\frac{2\pi l}{R_{lt}}\right) dl. \tag{18}$$

The aggregate value of land wealth, denoted LW_t, is given by the integral

$$LW_t = \int_0^{l_{\max}} p_{lt}^R\,(2\pi l)\,dl.$$

It will prove invaluable later when we discuss the simulations to have an index of the average house price at time t, which we denote p_t^{House}. We define this index as the total value of all land and residential structures divided by the stock of housing, that is

$$p_t^{House} = \frac{(LW_t + B_t)}{S_t}. \tag{19}$$

This is the total value of the housing stock divided by the quantum of housing.

1.6. *Equilibrium*

For an equilibrium we need a path for prices r_t, w_t and p_{0t}^S such that the goods, labour and housing markets clear. Rental prices, p_{lt}^S, at all locations are described in terms of the price at the CBD, p_{0t}^S, in equation (4). The price of land, p_{lt}^R, is also driven off rental prices, p_{0t}^S, and is

[8] The integrand in the final expression can also be written as $\frac{dS_{lt}}{dR_{lt}}\frac{dR_{lt}}{dl} = \frac{dS_{lt}}{dl}$ which perhaps makes clearer that this is the supply of housing services over the economy.

the present discounted value of future land rents as given in equation (15). All dynasties have an initial endowment of wealth equal to the initial land value LW_0 (equation (15)) plus capital stock, $K_0 + B_0$. That is $W_0 = LW_0 + K_0 + B_0$.

First, as labour is supplied inelastically to the production sector, the labour market clears if

$$w_t = (1-\alpha) A^{1/(1-\alpha)} \left(\frac{\alpha}{r_t + \delta_K} \right)^{\alpha/(1-\alpha)}.$$

The goods and housing markets clear if the following conditions are satisfied for all t.

(1) Given prices and an initial allocation of wealth, dynasties choose their consumption of the composite good, Q_t, so as to maximise their welfare (1) subject to the budget constraint (5). The necessary conditions for optimality are the intertemporal efficiency condition

$$\gamma \frac{\dot{Q}_t}{Q_t} = r_t - \theta - \frac{\dot{p}_{0t}}{p_{0t}}, \tag{20}$$

where the composite price at the centre is defined $p_{0t} = \left(a^\rho + \left(p_{0t}^S \right)^{1-\rho} (1-a)^\rho \right)^{1/(1-\rho)}$, and the transversality condition

$$\lim_{t \to \infty} W_t e^{-\int_0^t r_\tau d\tau} = 0,$$

where wealth evolves as $\dot{W}_t = r_t W_t + w_t - p_{0t} Q_t$.

(2) The path for the total capital stock, $K_t + B_t$, is described by the dynamic motion equations (6), (7) and (8) where C_t is given from equations (2) and (17). In equilibrium, the net returns to capital, K_t, and residential buildings, B_t, satisfy the efficiency conditions

$$r_t = \frac{\partial F}{\partial K_t} - \delta^K = \alpha A \left(\frac{L_t e^{gt}}{K_t} \right)^{1-\alpha} - \delta^K, \tag{21}$$

$$r_t = \left(p_{lt} \frac{\partial H (B_{lt}, R_{lt})}{\partial B_{lt}} - \delta_B \right) \quad \text{for all } l, \tag{22}$$

and the aggregation equation (18). The market clearing condition is that the supply of housing services is equal to the demand for housing services (all dynasties are housed)

$$1 = D_t = \int_0^{l_{t,\mathrm{Edge}}} \left(\frac{2\pi l}{R_{i(l)t}} \right) dl, \tag{23}$$

where dynastic density, $1/R_{i(l),t} = (S_{lt}/R_{lt})/S_{i(l),t}$, is a function of the composite good, Q_t and prices as given by equations (2) and (11).

1.7. *BGP*

Effective labour supply grows at the sum of the rate of productivity plus population growth, $g + m$. If the travel tax, λ_t, falls at half this rate, $(g+m)/2$, then, as long as the urban expansion does not approach the edge of the country, $l_{t,\mathrm{Edge}} \ll l_{\max}$, the economy will tend towards a BGP where all economic aggregate quantities grow at the rate $g + m$ and the average prices of land and housing are constant. Our model, therefore, admits a BGP, even though one of the factors is land and is in fixed supply. (A formal proof is in Appendix B.)

The intuition for this result[9] is this: at unchanged prices for consumer goods and for land and housing, the demand for all commodities rises in line with the sum of population growth and productivity. This implies that with no change in relative prices the demand for land for housing rises at this rate of $g + m$. While the total stock of land is fixed the amount of land that is within a given cost of commuting to urban centres grows if transport efficiency (commuting speed) rises—provided that there is some undeveloped land still available at the urban periphery. The proportional increase in the radius of a city consistent with no rise in travel costs is the proportional rate of change in travel speed (which is the proportional fall in λ). Since land used is proportional to the city area, and the city area is proportional to the square of its radius, we obtain that at unchanged prices the city area can grow at twice the rate of improvements in travel speeds. If that area is to grow in line with demand for land and for housing we therefore require that travel improvements rise at a rate of $(g + m)/2$.

1.8. *Alternative Patterns of Urban Development*

The conditions for a BGP are exactly the same if there are multiple urban centres each of which spreads out in expanding circle as the economy grows at a rate of $(g + m)/2$. But the conditions for balanced growth would change if cities start to overlap; then there is a slightly different condition for balanced growth. If, for example, there were two main centres of activity within a country then the rate of expansion of the area around each centre where it was feasible for people to commute from on a balanced path would grow at twice the rate of travel improvements until those two areas overlapped. From then on the pace of expansion of the feasible area would—for a given speed of travel improvement—initially slow down. Thus the condition for a BGP would be that the pace of travel improvements would need to rise for a period once the cities overlap to offset that. Ultimately the cities would once again increasingly resemble a single circular city so the balanced growth condition would revert towards its original level.

Circular development for non-overlapping cities is, nonetheless, restrictive. But it has an economic logic—circular development is efficient because it maximises the expansion of the feasible commuting zone for a given increase in travel speeds.

2. Calibration

We outline how we set key parameters. For some parameters we can take guidance from values that are implied by steady state growth paths. As we argue below, the evolution of productivity, population and transport costs mean that for many developed economies the condition for steady state growth (λ falls at the rate $(g + m)/2$) may have approximately held for roughly the 100 years 1860–1960 and so key aggregate ratios for that period help in calibration.

λ: we assume that the cost of distance per mile (λ) is proportional to the inverse of the average speed of commuting. In the Appendix we present evidence that travel speeds were most rapid between the mid-nineteenth century and the World War II; they then slowed until the 1970s and—at best—have been static since then. Based on that evidence we assume λ falls by 1.5% a year between 1870 and 1945, then falls by 0.8% a year until 1980 and is then constant.

[9] We thank a referee for encouraging us to explain the intuition in this way.

m: US population rose at an average annual rate of around 1.2% in the 100 years from 1911 to 2011. It has grown slightly more slowly in the 50 years to 2011 at rate of around 1.1% a year. We set population growth at 1.0% a year.

g: We set g at 0.02 based on historical long-run growth in productivity in many developed economies of around 2%.

$\boldsymbol{\delta_K}$: On a steady state growth path $\delta_K = (I^k/K) - (m + g)$. Using US data on the average ratio of non-residential capital investment to the non-residential capital stock since 1929, and using the values of $m + g$ as above, implies a depreciation rate of just above 6%. For the USA, Davis and Heathcote (2005) use a quarterly value for depreciation of business capital of 0.0136 (annual of around 5.4%). Kiyotaki *et al.* (2011) use 10%. We set depreciation at 7% a year.

$\boldsymbol{\delta_B}$: As with non-residential capital, on a steady state growth path the depreciation of residential capital is given by $\delta_B = (I^B/B) - (m + g)$. (ed.) Using US data on the average ratio of residential capital investment to the residential capital stock since 1929, and using the values of $m + g$ as above, implies a depreciation rate of only around 1.25% a year. This seems slightly lower than estimates based on the difference between gross and net US residential investment which gives a figure near 2%. We set depreciation on residential structures at 2% a year.

$\boldsymbol{\gamma}$: There is much evidence that the degree of intertemporal substitutability is less than 1. Hall (1988) estimated it was close to zero. Subsequent work suggests a significantly higher value, but still less than unity (see Ogaki and Reinhart, 1998; Vissing-Jørgensen, 2002). We set the intertemporal elasticity to 0.67 which implies $\gamma = 1.5$.

$\boldsymbol{\theta}$: On a steady state path $r = \theta + \gamma(m + g)$. Given the values used for γ, m, g we can use this relation to gauge a plausible value of θ conditional on an assumed value for the steady state rate of return. In our model all assets (land, non-residential capital, structures, housing) generate the same return. In practice assets obviously do not generate the same average returns. A recent paper Jordà *et al.* (2019) provides data on the real returns on a range of assets (including equities, bonds and housing) over the period 1870–2015 for 16 advanced economies. Returns on equities and housing look similar and average about 7% a year—though they are a little lower pre-1950. Bonds generate a lower real return which averages about 2.5% over the whole sample. The equally weighted average of the three asset classes is close to 6%. A figure of 6%–7% seems reasonable for the past average return on real assets. If we assume the steady state real rate of return is around 6.5% then based on the values for γ, g, m above (respectively 1.5, 0.02 and 0.01) the implied value of θ is around 0.02. That is the value we take for the rate of time preference.

$\boldsymbol{\alpha}$: We set α (the share parameter in the production function) to the typical share of capital in private domestic value added in developed economies in recent years. This figure is around 0.3, Rognlie (2016).

$\boldsymbol{\varepsilon}$: Muth (1971) estimates the elasticity of substitution between land and structures in producing housing at 0.5; later work finds a slightly higher level, but well under 1. Thorsnes (1997) puts estimates in the range 0.5 to 1. Ahlfeldt and McMillen (2014) suggest it might be a bit under 1. Kiyotaki *et al.* (2011) constrain it to 1 in their calibrated model. But the weight of evidence is for a number under 1. For our base case we use a value of 0.5. We also consider higher values.

$\boldsymbol{\rho}$: There are many estimates from the empirical literature on housing of the elasticity of substitution between housing and consumption in utility. Ermisch *et al.* (1996) summarised that literature and put the absolute value at between 0.5 and 0.8; Rognlie (2016) uses a range of 0.4 to 0.8. Kiyotaki *et al.* (2011) constrain it to 1 in their calibrated model. Van Nieuwerburgh and Weill (2010) use 0.5 for the price elasticity of demand for housing, basing their choice on micro studies. Albouy *et al.* (2014; 2016) find strong US evidence for a value of 2/3. For our base case we use a value of 0.6; we also consider higher values.

In assuming that λ falls by 1.5% a year between 1870 and 1945 we have set the improvement in transport efficiency between the mid-nineteenth and mid-twentieth century equal to one-half the sum of the rate of productivity growth (assumed to be 2%) and population growth (assumed to be 1%). Those productivity and population figures are plausible for the USA. It is also plausible that in the land-rich USA there remains undeveloped land suitable for housing at the periphery of many existing residential areas. This is a condition for balanced growth—we must not have reached the 'edge' of the country. (It is a condition less likely to hold for many countries in (relatively crowded) Europe and in Japan—a point we shall return to below.) So it makes sense to judge the plausibility of the overall calibration by whether it generates ratios over a balanced growth period that are in line with US aggregate ratios from the mid-twentieth century.[10] Given this criterion we set the initial level of transport costs ($\lambda_{1870} = 0.36$—the final parameter to be set) to generate a level of land values relative to GDP along a BGP that roughly matches the US value of residential land to GDP over the decade 1950–60.

We solve the model under the assumption that agents are rational and forward-looking. Since there are no stochastic elements this means we are looking for a perfect foresight path that satisfies all the equilibrium conditions including the transversality conditions expressed over an infinite horizon. In practice, we need to solve the model over a finite horizon subject to a terminal condition. We therefore need to appeal to a Turnpike theorem of the early or second kind,[11] McKenzie (1976). Under relatively weak conditions (that are all satisfied by this model)[12] the optimal path over the finite horizon will 'hug' the infinite horizon optimal path before 'turning off' to reach the imposed terminal condition. In presenting the results we need to therefore check that the path has not turned off. We do this by solving the model over 450 and 650 years, check the solution for the first 250 years is the same (the difference is less than 10–6) and present the solution for this period only. (More details of how we solve the model based on this calibration are in Appendix C.)

Because households are forward-looking this means that even when the condition for a BGP still hold (that is the pace of technology improvements in transport is one-half the growth in GDP) house prices are not constant. This is because the perception that prices will rise in the future once transport improvements slow boosts demand in anticipation of some capital gains. That raises demand ahead of the actual slowing in transport productivity and takes demand growth ahead of the capacity to create more housing at a constant price. For the model calibration described above we find that this raises prices somewhat a few decades before the start of the slowing in travel speeds, but the deviation from constant prices more than ten years prior to the conditions for balanced growth no longer holding are relatively small. We put the period when travel improvements slow as around 1970 and we find that deviations from a BGP (when house prices are flat) more than ten years before that are small. That is why we chose to calibrate some parameters so that the BGP of the model matches data from the USA in the decade 1950–60, a period near the end of the era when the conditions in the model for balanced growth hold but

[10] There are are two share parameters that, in conjunction with other parameters set above, we set to try to match key features of the mid-twentieth-century US data. These parameters are: a: the weight on consumer goods (relative to housing consumption) in utility; b: the weight of buildings in production of housing. We set $a = 0.85$ and $b = 0.78$.

[11] Sometimes referred to as the classical conception due to its origins in the original debate between Samuelson *et al.* (2012).

[12] Our set-up satisfies the weak assumptions of McKenzie (his Theorem 1); our utility function is strictly concave with the existence of an interior path connecting the initial and terminal capital stock conditions.

Table 1. *Comparison of Model Generated and Observed Ratios.*

	US average (%) 1950–60	Ratios along balanced path (%)
Return on capital	6.5	6.5
Housing/total consumption	13.8	14.1
Total consumption/GDP	77	85
Gross profit share	35.4	37.8
Net profit share	26.8	26.2
Buildings (B)/GDP	102	96
Residential land to GDP	82	87
Capital (K)/GDP	180	197
Total wealth/GDP	364	381
Total wealth/NDP	413	452

Source. US Bureau of Economic Analysis; Board of Governors of the Federal Reserve System. Return on capital is profits net of depreciation plus consumption of housing services relative to the sum of the value of residential and non-residential capital. Buildings are values of residential fixed capital. Residential land is estimated by subtracting the value of residential buildings from the reported value of residential real estate. Total wealth is the sum of the values of land, residential structures and non-residential fixed assets.

about ten years before transport improvements slowed significantly. That decade was a period when there was a significant movement in the USA into suburban areas, a phenomenon consistent in our model with being on a BGP with house prices flat as improving travel speeds expanded the areas around cities where it was attractive to live.

3. Results

3.1. *Balanced Growth and Beyond*

Table 1 shows what the model generates for key ratios on a BGP and compares them with ratios for the USA averaged over the period 1950–60 (the decade at the end of the period when our model suggests the USA was on a BGP).

The table shows that the BGP of the calibrated model matches the US data (on aggregate real economy ratios) from the mid-twentieth century quite closely. This is not guaranteed because the number of free parameters, that is parameters not set in line with evidence from other econometric studies but explicitly set to help match the historical US ratios shown in the table, is less than the ratios we are trying to match.[13] Nonetheless Table 1 is more a validation of the calibration than of the model itself. So first we consider whether the model generates results consistent with aspects of the history of residential development and which the calibration was not designed to match. We consider the case where travel improvements slow over time but where undeveloped land for residential development at the periphery remains available. We think of this set of simulations as applying to a land-rich developed economy—so as in Table 1 we compare model properties with what has happened in the USA. When undeveloped land is in much tighter supply things look different—we think of the UK as the quintessential rich economy where this applies and we consider what differences that makes in due course.

Panel A of Figure 2 shows the model generated path for an index of real house prices, p_t^{House}, defined in equation (19). The numbers here are population weighted averages across locations at each point in time normalised to 1 in 1870. We set $\varepsilon = 0.5$ and $\rho = 0.6$ and assume land at the

[13] There are only three completely free parameters: a: the weight on consumer goods (relative to housing consumption) in utility; b: the weight of buildings in production of housing; λ_{1870} the initial level of the distance tax.

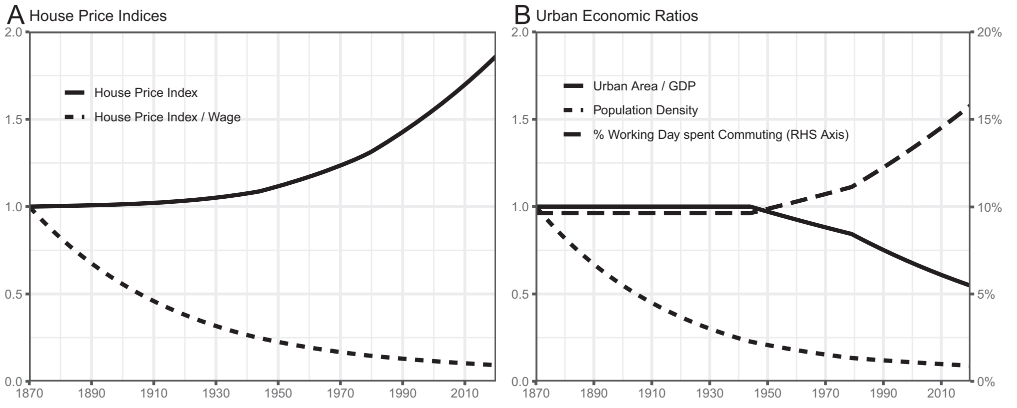

Fig. 2. *Simulated Path for House Prices and Other Economic Ratios 1870–2020.*

periphery of existing residential developments consistently remains available. Other aspects of the spatial distribution of the population are illustrated in Panel B. Figure 3 shows how housing conditions vary by location at different points in time. The horizontal axis here is distance from the centre measured as a proportion of a fixed distance which is the periphery that residential development had reached in the simulations by 2020.

Figures 2 and 3 illustrate what the model implies about the path the land-rich economy follows over the past 150 years:

(i) Real house prices are flat over the period when transport speeds were running at around 1.5% a year, they then rise gradually up to 1970 and rise faster after that when travel speed improvements peter out and finally fall.
(ii) As transport costs decline rapidly through the late nineteenth century and up until about the middle of the twentieth century, people commute greater distances and, on average, live further from the centre.
(iii) The average amount of time people spend commuting is constant at around 10% of a typical work week until transport improvements slow and then starts to gently rise.
(iv) The density of population at urban centres declines over time.
(v) The decline in house prices and rents as you move further urban centres become less steep over time.
(vi) The ratio between areas where people live and GDP is constant over a long period but in recent decades starts to fall.

How do these predictions match up to the historical record?

The model predicts that when there remains land that can be developed there is very little change in real house prices (or in the real level of rents) between 1870 and the middle of the twentieth century. Between around 1945 and 2020 the average house price then rises by around 70%. Based on the findings of Knoll *et al.* (2017) this path of house prices matches the overall trajectory of average prices across the USA. It is also consistent with Shiller's US house price data which show the real house price index was very nearly the same in 1945 as it had been in 1890, but that since 1945 (up to mid 2019) real prices are up by just over 70%, with most of the

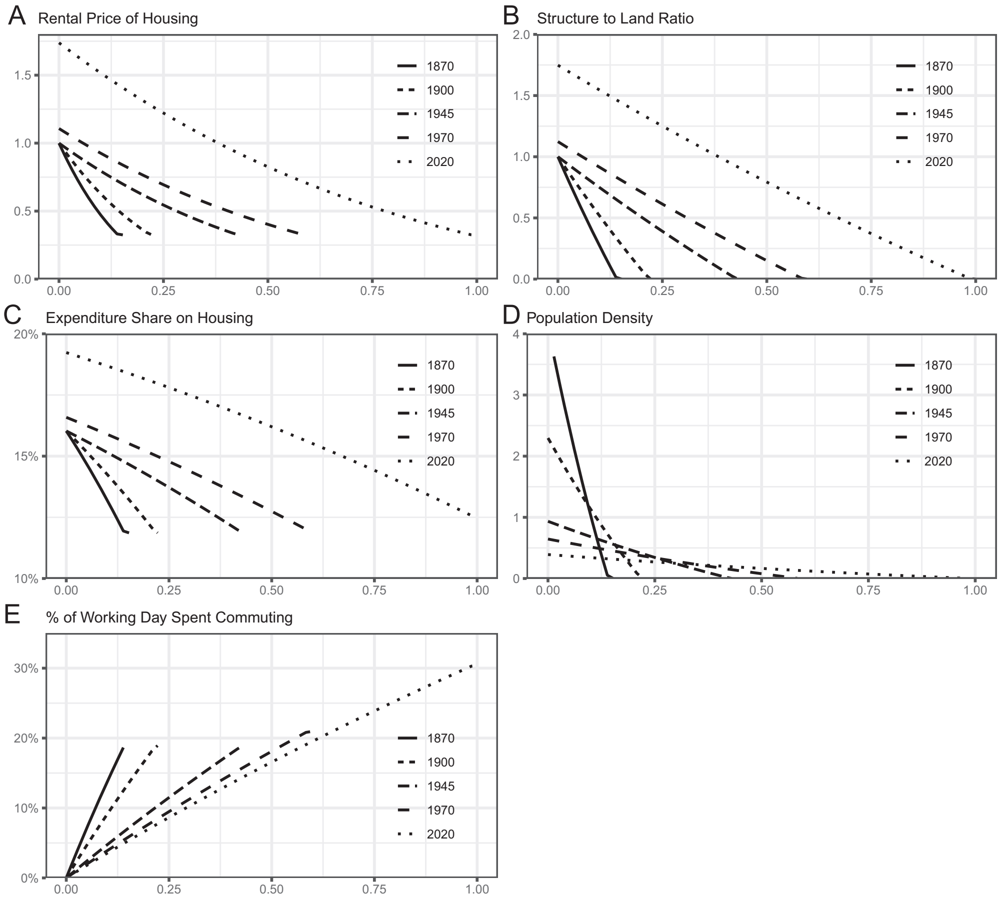

Fig. 3. *The Distribution of Activity across Time and Location.*
Notes: In all panels the *x*-axis is distance from the CBD, normalised by the distance to the periphery of urban development in 2020. Panel A plots the rental rates normalised so that the rental price at the centre in 1870 is 1. Panel B plots the structures to land ratios normalised so that ratio at the centre in 1870 is 1. Panel C is the % of household expenditure on housing. Panel D is the dynastic population density. Panel E is commuting time as a fraction of working time.

acceleration coming in the period since the mid-1980s. But the overall pattern of house prices shown in Figure 2 is not right for the more densely populated European countries and Japan where since 1945 average real house prices more than trebled. Figure 2 shows that house prices on average do not rise by as much as wages—though the fall in the ratio of prices to wages flattens off in recent decades; this also matches the USA but not more densely populated countries.

The calibrated version of the model implies that in the USA rental yields should have been fairly flat over the late nineteenth century and for the first half of the twentieth century until the condition for balanced growth no longer holds. Recent estimates of long-run trends in US rental yields in Jordà *et al.* (2019) show that rental yields did seem steady until well into the second half of the twentieth century. The model generates a gradual increase in house prices relative to rents once we move off BGPs and such a rise has been seen in the USA (and the UK). The Jordà *et al.* data put the average US rental yield in the period from 1890 to 1945 at around 5.6%; that yields falls to around 4.8% by 2016. Our model puts the rent to house price ratio at 6% on the steady state path; that ratio falls once house prices begin to rise. The reason the model generates a falling rental yield is that the slow and gradual decline in travel improvements is assumed to be recognised by agents who expect rising house prices. This reduces the required rental yield on housing. In our simulations for the USA the rental yield falls to around 4.6% by 2016. But in the USA house prices relative to rents did rise very fast in the period just before the 2007–8 crash and then fall back sharply. Our model focuses on long-run trends and is ill-suited to account for such cyclical variability.

The predictions that average commuting distances have risen significantly over time and that price gradients (the fall in rents and house prices as distance from the centre rises) have become less steep are strongly confirmed by the historical evidence. We noted above the steady upward trajectory of commuting distances over time. The decline in the house price–distance gradient over time has been noted many times in the literature and goes back at least to Mills (1972) who noted evidence of flattening in the gradient in the USA which began in the nineteenth century. There is evidence of flattening in land price gradients for numerous cities, including: Chicago (McMillen, 1996); Berlin (Ahlfeldt and Wendland, 2011); Cleveland (Smith, 2003); New York (Atack and Margo, 1998); Sydney (Abelson, 1997).

The model predicts that the developed areas around cities should have consistently expanded (as has population) but that the density of population at the centre declines and density become more even across the country. Data from the Atlas of Urban Expansion (2016, vol 1 available at www.atlasofurbanexpansion.org) shows that density in the largest cities in developed economies in recent decades has consistently fallen, continuing a trend that has existed for much of the twentieth century.

The model implies that travel times should have been steady at around 10% of the average working day until a few decades ago implying travel times of around an hour a day. There is a good deal of evidence that average time spent travelling is surprisingly constant over time and across countries at about one hour a day—around 10% of what were average working days of around ten hours before the World War II. Levinson and Kumar (1994), Ausubel *et al.* (1998), Schafer (2000), Schafer and Victor (2000) all present evidence that, a least up to the 1990s, there has been near constancy of travel times, supporting a hypothesis originally put forward by Zahavi (Zahavi and Talvitie, 1980). Figure 1 of Schafer and Victor (2000) is particularly striking, revealing a clustering of average daily travel times at around one hour for countries of very different standards of living and across different periods up to the early 1990s. But the model predicts that this time will have risen gently in the period after travel improvements slow—a clear feature of recent US and UK data.[14]

The model has the property that on a BGP urban areas should grow at the rate of their urban GDP. That growth in output will be faster than population growth if labour productivity is rising and so the rise in the urban footprint on a BGP will tend to be greater than population growth so density should gradually fall. The model predicts that slowing transport improvements which take one off a BGP would mean that the ratio of urban areas to their GDP would start to fall. Table 2 uses data on the size of New York and Chicago (but also Paris and London) and an estimate of those cities' GDPs over time to assess what has happened to the city output–area

[14] See 'Home-to-Work Commute Profile, United States, 1977–2017', U.S. Department of Transportation, Bureau of Transportation Statistics. See also Figure 29 of 'Commuting Trends in England 1988–2015', UK Department of Transport (2016).

Table 2. *Urbanised Area (Hectares) Divided by City GDP ($million).*

Year	City pop. (million)	Urban area (ha) 000s	Density (pers./ha)	GDP per capita (1990 $) 000s	Hectares/city GDP	Year	City pop. (million)	Urban area (ha) 000s	Density (pers./ha)	GDP per capita (1990 $) 000s	Hectares/city GDP
			Paris						London		
1800	0.58	1.2	497	1.09	1.84	**1800**	0.96	3.6	265	1.62	2.33
1832	0.89	1.8	488	1.31	1.56	**1830**	1.66	5.1	327	1.75	1.75
						1845	2.10	7.8	272	2.07	1.78
1855	1.16	3.4	339	1.62	1.82	**1860**	2.76	10.8	256	2.83	1.38
1880	2.27	7.8	292	2.12	1.62	**1880**	3.76	16.2	233	3.48	1.24
1900	2.71	22.0	123	2.88	2.82	**1914**	7.32	37.2	197	4.93	1.03
1928	2.88	28.7	100	4.43	2.25	**1929**	8.05	68.0	118	5.50	1.53
1955	6.28	63.8	98	6.20	1.64	**1954**	8.29	107.5	77	7.62	1.70
1979	8.79	93.4	94	14.63	0.73	**1978**	7.63	161.3	47	12.83	1.65
1987	9.28	164.2	57	16.16	1.10	**1989**	9.93	202.2	49	16.41	1.24
2000	10.01	231.4	43	20.42	1.13	**2000**	9.74	243.0	40	20.35	1.23
2014	11.11	277.8	40	22.22	1.12	**2013**	11.20	250.8	45	23.74	0.94
			New York						Chicago		
1800	0.06	0.3	207	1.26	3.84	**1871**	0.31	8.3	38	2.50	10.52
1840	0.31	1.4	227	1.59	2.78	**1893**	1.25	29.7	42	3.48	6.80
1850	0.52	1.8	280	1.81	1.98	**1915**	2.43	39.1	62	4.86	3.31
1900	4.27					**1945**	4.12	90.0	46	11.71	1.87
1950	12.60					**1969**	7.01	249.2	28	15.18	2.34
1989	16.24	688.9	24	22.85	1.86	**1989**	7.56	451.8	17	23.06	2.59
2000	17.96	922.7	19	28.47	1.81	**2001**	8.51	629.9	14	28.41	2.61
2013	18.41	951.1	19	31.18	1.66	**2014**	8.91	700.8	13	31.18	2.52

Source. Atlas of Urban Expansion; Maddison Project Database, version 2018. The figures for urban area and density after 1980 refer to those cited as ‘total urban extent’ in Angel *et al.* (2016). The figures prior to 1980 for New York were accessed directly from http://hdl.handle.net/2451/33846, with those for the other cities kindly provided in separate communication with the authors. GDP per capita figures are from Bolt *et al.* (2018) and are in 1990 international Geary–Khamis dollars.

ratio. The GDP estimates are crude: population of the cities multiplied by average GDP per capita for the country. Effectively we assume a constant ratio of per capita incomes in the capital to the rest of the country. The built-up urban areas of the cities over the past 200 or so years are more reliable, as are population estimates.

The table shows an enormous expansion of the urban footprint for all these cities over time. City output and population also increase massively. Once we get past the initial surge in population as these places begin to emerge as mega cities, we see a steady and consistent decline in population density over time—which is what the model predicts on the BGP. The table also shows that the ratio of urban area to urban GDP is, relative to the enormous variability in city area and output, very stable. Once again, after the initial surge in population, the ratio of area to GDP is fairly flat. In Paris and in London that ratio was about the same in 1950 as it had been in 1850. In New York the ratio is also flat between the middle of the nineteenth and twentieth centuries. In Chicago the ratio seems to stabilise a bit later.

There is also some evidence in the table that the area–GDP ratio has shown a more recent tendency to decline, which is what the model predicts will happen once we are off the BGP if transport improvements slow down.

3.2. *The USA and UK Compared: 1945–2020*

The model seems to generate patterns of residential development that fit many of the historical trends, both in the USA and across many other developed economies. But for house prices while the simulations match the long-run pattern in the USA very well, they substantially under-predict price rises over the past 50 years for most other developed countries. Now we focus specifically on house prices and on whether the model helps understand the similarities and differences between trends over the long term in home prices in the USA and the UK. The USA and UK share a similar history in terms of changes in transport technology, infrastructure and travel patterns: they developed railways at much the same time; saw horse-drawn vehicles overtaken by cars—and then cars overtake trains—at similar times. They have also seen slowing—and then some reversal—in the rise in commuting speeds at about the same times. But one major difference is that the area around many urban centres—and the scope to develop new urban centres—is now much more constrained in the densely populated UK than in the USA. We explore what the model has to say about how these factors may account for the history of prices in the two countries.

Areas that can be developed around many (though certainly not all) US cities are more plentiful than is the case in much of Europe, and within Europe the UK is the most densely populated large economy. Land area relative to population in the USA at around 2015 was about eight times as great as in the UK. Figure 4 shows two trajectories for house prices (in real terms and relative to wages): one in which residential development does not reach the edge of the country over the whole horizon ('USA') and one in which it does by 1990 ('UK').

The UK path for house prices in the figure is from a simulation with the same calibration as for the USA (with slowing transport improvements after 1945) except that the area of the economy relative to population is set to be much lower. In this simulation the available land area of the economy is such that by 1990 there is for the first time some (very low density) residential development at the edge of the country.

In the period since the end of the World War II until 2020 the simulations show both rents and real house prices to have increased hugely more in the land-poor UK than in the US simulation

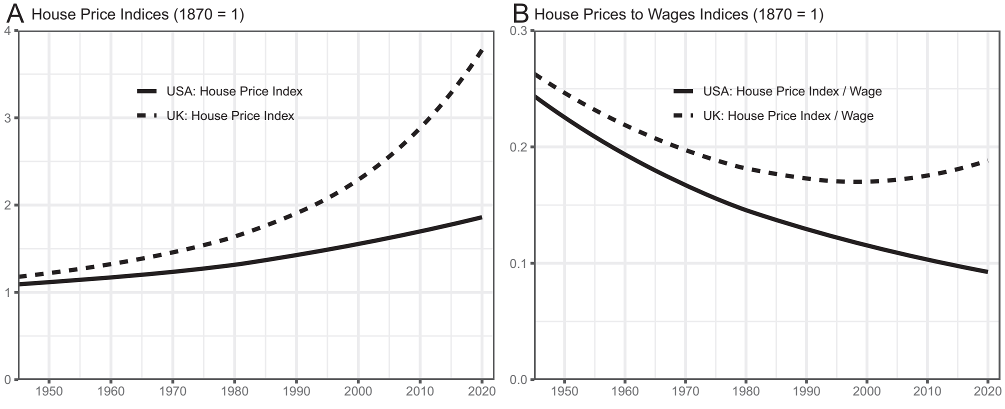

Fig. 4. *Projected Path for House Prices and House Price to Wages 1950–2020.*

Table 3. *US House Prices: Percentage Change in Average House Prices by City.*

	New York	Boston	San Francisco	Dallas	Atlanta	Chicago
1991–2020	174	249	383	n/a	123	104
2000–2020	104	126	169	93	56	44

Source. S&P/Case-Shiller Home Price Indices: retrieved from Federal Reserve Bank of St. Louis.

where only the slowing of transport improvements took us away from the BGP: house prices by 2020 have more than tripled relative to the level in 1945 when we have both the petering out of transport improvements *and* the ending of freely available undeveloped land. That is much more in line with the UK experience than a simulation with far more land relative to population. A continuation of these conditions would see house prices continue to rise far faster in the UK than in the USA and very likely bring further increases in the UK house price to income ratio.

In thinking about national house price growth over time it seems reasonable to think that on average land availability for residential development around existing cities is more plentiful in the USA than in the UK. But there are US cities where available land around the edges is constrained. Such differences are very marked across some major US cities: around Atlanta and Dallas there is greater scope to spread the urban fringe than in San Francisco or Boston. Our model predicts that once transport improvements fall away those US cities that had more constraints on spreading should have seen greater price rises than cities where this was less of a problem. Table 3 shows that prices rises over the past few decades seem consistent with this. Price rises for New York, Boston and San Francisco over the past 30 years have, on average, been twice as great as for Atlanta, Chicago and Dallas.

3.3. *Zoning Restrictions and Welfare*

Restrictions on the ability to build residential structures at various locations have played no role thus far. But there is a good deal of evidence that such restrictions do play a role; they might even be a much more significant factor in explaining rising house prices in recent decades than a slowing in transport improvements and more significant than countries running out of undeveloped land (see for example Quigley and Raphael, 2005; Jaccard, 2011).

We explore this issue by introducing restrictions on building into our model. We impose a limit on the ratio of residential building-to-land area at different locations—that is an upper limit on B_l/R_l.

We consider uniform limits on B_l/R_l that apply at all locations. We also consider an alternative—more realistic—restriction that from a given date (which we take to be 1970) then at all locations beyond a certain distance $v = 0.2$[15] from the centre the building density cannot rise beyond the 1970 value for that density at distance $v = 0.2$. Thus for all $l > v$, $B_l/R_l \leq (B_v/R_v)_{1970}$. We freeze all the building densities in the area closer than $v = 0.2$ to the centre at their 1970 values. This second set of restrictions is a form of zoning that allows existing structures within the v radius to continue standing (but not to be added to) and puts a limit at all points beyond distance $v = 0.2$ that development can never be more intense than it was at that location in 1970.

Tables 4 and 5 show the results. For the fixed limit on B/R we consider limits at 2 and 3. For the zoning restriction introduced in 1970 we consider limits that are tighter again. We focus on those results from the tighter restrictions to gauge how important zoning restrictions can be.

Without zoning restrictions and no shortage of land the only reason for land and house prices to rise is due to slowing transport improvements. The tables show that if that was all that happened the impact on welfare and house prices (relative to the BGP) are very substantial.

Focusing first on welfare, we calculate the compensating variation in incomes (or expenditure) that would leave utility unchanged when transport improvements slow down relative to a world where there was no slowdown so the economy remained on a BGP. The ending of transport improvements we assume began in 1980 and by 2020 means that incomes would have need to be almost 8% higher. By 2050 that becomes over 14% and by 2070 the compensating variation is close to 20%. Adding in zoning restrictions that start in 1970, and which by 2020 are binding across much of the economy, does not reduce welfare further by a large magnitude. The compensating income effects rise from just under 8% to 9% in 2020, from 14.3% to 17.4% in 2050 and from 20% to 24.7% in 2070. So the welfare effects of zoning are ultimately around one-quarter of the effects of slowing transport improvements. This is not because the building restrictions are hardly binding. Without restriction building density (B/R) near the central location in 2070 would be almost three times as great.

The impact on house prices reveals something similar: zoning restrictions are much less powerful than are the effects of the assumed slowdown in transport improvements. In 2020 slowing travel improvements raise average house prices relative to balanced growth by over 80%; by 2050 prices are higher by a factor of 2.5 and by 2070 by a factor of just over 3. The impact of zoning restrictions on top of that is by no means trivial—by 2070 prices would be up by a factor of 3.6 rather than just over 3. But the effects of zoning shows up as consistently being roughly one-fifth that of slowing transport improvements.

This finding that travel costs are in the long run more important than building restrictions is one we think plausible. There is a sense in which transport technology is ultimately a more powerful and fundamental force than planning restrictions in and around cities. If people could safely and reliably travel 200 miles to the centre of a city in 30 minutes then planning restrictions on residential developments in a 10-mile radius about the centre of that city would not be very important. To take just one example. London has had a green belt around it for many decades and that has restricted building new houses. But even with such a green belt, had transport improvements continued at the pace of the nineteenth and early twentieth

[15] Distance, l, was normalised so that in 2020 the distance to the periphery of the developed area is 1.

Table 4. *Impact of House Building Constraints on Welfare and House Prices; All Parameters Set to Base Values.*

	Required compensating expenditure					Average house price					
	Diminishing growth in transport speeds						Diminishing growth in transport speeds				
		and						and			
Year	Only	Running out of land	$B_l/R_l \leq 3$	$B_l/R_l \leq 2$	$B_l/R_l \leq (B/R)_{1970}$	Balanced Growth	Only	Running out of land	$B_l/R_l \leq 3$	$B_l/R_l \leq 2$	$B_l/R_l \leq (B/R)_{1970}$
1950	0.2%	0.2%	0.2%	0.2%	0.2%	1.00	1.12	1.23	1.13	1.14	1.17
1980	1.8%	1.8%	1.8%	1.9%	1.7%	1.00	1.31	1.64	1.33	1.37	1.44
2000	4.5%	4.5%	4.5%	4.7%	4.9%	1.00	1.55	2.28	1.59	1.63	1.74
2020	7.9%	8.2%	7.9%	8.3%	9.0%	1.00	1.86	3.72	1.92	1.98	2.11
2050	14.3%	18.0%	14.6%	15.7%	17.4%	1.00	2.48	10.06	2.60	2.69	2.89
2070	19.6%	29.4%	20.4%	22.1%	24.7%	1.00	3.05	22.05	3.21	3.34	3.60

Notes: The left table records the compensating expenditures with respect to the BGP (% increase in household expenditure necessary for households to achieve the same utility as on the BGP). For all other columns the growth rate in transport speeds fall after 1945 from 1.5% to 0.8% and then 0% after 1980. For the first column this is the only difference. The other columns have an additional constraint. In the second column, the urban area hits the 'edge' in 1990; in the third the structure–land ratio is constrained to be 3 or below; in the fourth this is constrained to be 2 or below; and in the fifth this ratio is constrained to be on or below its level in 1970 out to a radius of 0.2 and no higher than this value further out. The right table records the average house price. Here there are six columns. The first records the average house price along the BGP (which is constant and normalised to 1) and the other five columns give the average house price with respect to the five scenarios described in the left table.

Table 5. *Impact of House Building Constraints on Welfare and House Prices; Parameters Set to Base Values, Except $\epsilon = 0.75$.*

	Required compensating expenditure						Average house price				
	Diminishing growth in transport speeds							Diminishing growth in transport speeds			
		and							and		
Year	Only	Running out of land	$B_l/R_l \leq 3$	$B_l/R_l \leq 2$	$B_l/R_l \leq (B/R)_{1970}$	Balanced Growth	Only	Running out of land	$B_l/R_l \leq 3$	$B_l/R_l \leq 2$	$B_l/R_l \leq (B/R)_{1970}$
1950	0.2%	0.2%	0.3%	0.5%	0.2%	1.00	1.08	1.10	1.12	1.15	1.18
1980	1.7%	1.7%	1.7%	2.1%	1.5%	1.00	1.23	1.29	1.33	1.38	1.50
2000	3.9%	4.0%	4.2%	4.8%	4.7%	1.00	1.43	1.56	1.59	1.65	1.81
2020	6.6%	6.7%	7.4%	8.5%	8.9%	1.00	1.67	1.97	1.91	2.00	2.20
2050	11.3%	11.9%	13.9%	15.8%	17.4%	1.00	2.14	3.14	2.57	2.72	3.00
2070	15.2%	16.5%	19.6%	22.2%	24.8%	1.00	2.56	4.60	3.17	3.38	3.74

Notes: The left table records the compensating expenditures with respect to the BGP; the right table records the average house price. The scenarios are identical to those in Table 4 and are described in the notes to that table.

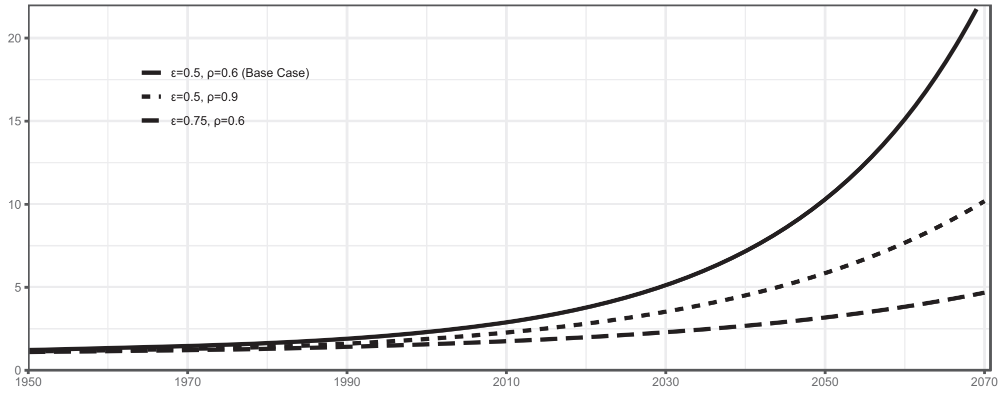

Fig. 5. *Projected Path for House Prices 1950–2070 Assuming Different Elasticities.*

centuries it would have allowed the population that could work in London to expand as it became feasible to live ever further beyond the green belt in areas where residential development was allowed.

The effects of restrictions on building residential structures are greater if land and structures are more substitutable: putting limits on structure is more onerous if, as transport improvements fall, you would want to substitute more towards structure and economise on land. Table 5 shows this. Here we assume substitutability between land and structure, ε, is 0.75 rather than the base case of 0.5. The zoning effect is about 50% as great as the travel effect; that effect was only around 20% as great as the travel effect when ε was 0.5. In the next section we explore further the impacts of different elasticities.

3.4. *The Future and Its Sensitivity to Key Elasticities*

We now use the model to assess how house prices might evolve over the next several decades under various assumptions about the technology of house building, consumer preferences and the evolution of travel speeds. We focus on a land-poor developed economy where transport improvements have petered out; the UK fits this description well.

With the base case parameters then over the next 50 years (2020–70) real housing costs (rents) are projected to rise fourfold (Figure 5). House prices rise by much more: prices rise sixfold between 2020 and 2070. Unlike in simulations where we only have a slowdown in travel improvements (but still undeveloped land is available), both rents and house prices are projected to rise consistently faster than average incomes.

The key factor here is that with completely undeveloped land assumed to be exhausted, with relatively low substitutability between land and structure (0.5), less than unit elasticity of substitution between housing and goods (0.6) and with no assumed further improvement in transport costs then housing costs rise at much faster rates than in the past. The scale of this effect, as we shall see, is highly sensitive to those factors.

We assess the sensitivity of that projection of future house prices to three factors: (i) substitutability between land and structure in creating housing; (ii) substitutability of housing for consumption goods in utility; (iii) variations in the path for transport costs (speed of commuting).

(i) Greater substitutability between structures and land: $\varepsilon = 0.75$

When $\varepsilon = 0.75$ the path of house prices over the future looks very different. Over 2020–70 house prices fall relative to incomes by about 15% when $\varepsilon = 0.75$ rather than rising strongly faster than incomes when $\varepsilon = 0.5$. If $\varepsilon = 0.99$ rental costs and house prices barely rise any faster than the cost of consumer goods and they decline markedly relative to earnings to less than half their 2020 value by 2070. Clearly there is enormous sensitivity in the path of housing costs and of house prices to even relatively small changes in the substitutability between land and structure in creating homes. That degree of substitutability is partly a matter of technology and partly a matter of preferences, reflecting the fact that ε effectively plays a dual role as summarising production possibilities but also trade-offs in creating utility by consuming different combinations of structure and land (or building and space). This point was recognised by McDonald (1981). He noted that a model with utility a function of consumption of goods and housing, then if housing is 'produced' via a CES function combining land and structures, it can be interpreted as a weakly separable utility function rather than a model of the production of housing. This implies that estimates of the elasticity of substitution, ε, can be thought of as estimates of a production parameter or as an estimate of a parameter of a weakly separable utility function. The production function of housing can as well be thought of as an aggregator reflecting tastes. Both tastes and production technology can change and so one should not think of ε as fixed. The degree to which one can substitute structure for land has probably changed significantly over time—new building techniques now make it possible to build 100-storey apartment blocks with very small footprints. That creates a different trade-off between use of land and structure in creating units of housing than was available when the Empire State building was constructed—a building that had a footprint that was enormously larger relative to its height than the sort of pencil-thin apartment blocks recently constructed in Manhattan. But whether having a living space a mile up in the air—potentially shrouded in cloud for some of the time and remote from life on the ground—is really a good substitute for a second-floor apartment is not a question of technology. So while ε may rise with technological improvements—and put some limit to the rise in housing costs in the face of fixed land area—that scope is strictly limited by preferences which might be completely insensitive to changing housing costs.

(ii) Preferences between consumer goods and housing: $\rho = 0.90$

A moderately higher degree of substitutability between housing and consumer goods (0.9 versus 0.6 in the base case) has a very substantial effect. Over a 50-year horizon the impact is almost as great as varying the elasticity of substitution between land and structures. With the higher price elasticity of demand for housing, house prices by 2070 would be one-half the level reached in the base case.

(iii) Transport costs

In the base case we assume that improvements in travel speed stopped in 1980 and do not change over the next 50 years. If instead travel times from 2020 once again start to fall at 0.8% a year—the same rate as assumed between 1945 and 1980—there is, not surprisingly, predicted to be a somewhat more equal density of population over the country.[16] There is also greater investment in structures and overall consumption of housing is greater. For the 'land-poor' simulations that we think reflect conditions in the UK, the effects are not enormous—even if we go back to transport efficiency rising at late nineteenth-century levels. But for 'land-rich' simulations more relevant for the USA if transport improvements were to resume the impacts

[16] Detailed results for simulations under various assumptions about the evolution of transport costs are available on request.

are greater. If such improvements once again occurred at the rate typical in the late nineteenth century—so that the conditions for balanced growth once again held—house prices and rents would once again become flat and so fall sharply over time relative to incomes.

The 'land-poor' simulations reveal that the impact of further falls in transport costs on the overall level of housing costs over the future is far smaller than its impact in the past. We ran a simulation assuming no transport improvements at all between 1870 and 1970; the level of real housing costs in 1970 was twice as high as in the base case which calibrated travel improvements to evidence on speed of moving passengers. Yet assuming continuing improvements in travel speeds from now on had a small effect on house prices 50, or even 100, years ahead in land-poor countries relative to the assumption of no further changes in passenger speeds. Why do transport cost improvements have a relatively small impact upon the trajectory of average housing costs in the future and yet have been very significant over the past? The main reason is that there are diminishing returns to travel improvements. In our model the travel improvements between 1870 and the World War II substantially increase the proportion of the economy where it is viable to live. By 1990 (for the 'land-poor' simulations using UK population–land ratios) almost all the country has some residential development, although density is still very much lower nearer the periphery than at the centre. Until this period there had effectively been a major expansion in the physical size of the country that is relevant for the day-to-day lives of the population. After that date further improvements have some impact on density differences between areas that have already been developed, but they do not increase the area that has some housing on it.

3.5. *Caveats and Limitations*

In assuming that dynasties are long-lived and are alike (except for where they chose to live) we give no useful role to mortgages. Essentially we have representative agents so there are not distinct groups alive at the same time some of whom want to save and some of whom want to borrow. Obviously such a model has no meaningful role for mortgages or indeed any debt. It is not that holding debt and financial assets could not be envisaged in the model—we could certainly allow households to have mortgages and save in financial assets that both pay the same interest rate. But in a closed-economy model with representative households the only equilibrium would be one where each household held an amount of interest bearing assets equal to its interest paying debt. Neither does the model make any meaningful distinction between owner occupation and renting; the cost of either form of tenure is the user cost of housing. So the model does not tie down the owner-occupation rate. We could interpret the outcomes as being ones where some households rent and simultaneously own shares in property owning companies that rent out property and pay the returns to shareholders. Ultimately the members of dynasties alive at any time own all the housing stock—and it does not matter whether it is held as owner occupied property or as claims on properties that are rented out. Members of those dynasties pass housing wealth down to later members of the same dynasty. This does reflect an important aspect of reality: people do pass on lots of wealth (both at death and in life) and housing is, for most families, the largest part of bequests. But we completely miss out on issues of inequality and the impact of credit restrictions by making assumptions that dynasties are long-lived and are all alike except for where they live.

We take population growth as exogenous but note that over the long term it is plausible that it adjusts to house prices—which is probably one factor behind the faster rate of population growth in land-rich USA where house price growth has been slower than in many other land-poor developed countries where prices have risen far more.

We have assumed that the reason for urbanisation has been benefits to being close to the centre of population and activity. But in the nineteenth century and before greater urbanisation was significantly driven by the decline in the need for so many people to work in agriculture. We have not modelled that. The move from an agrarian economy is crucial in the nineteenth century and will have affected the changing pattern of density back then; but it will not have been very significant since the World War II and it is in this more recent period that transport improvements have fallen off, undeveloped land in many countries has becomes scarcer and when the big rises in house prices come in most economies. It is those factors which play a central role in our analysis.

4. Conclusions

This paper develops a model of the changes, over time and across locations, in housing and of housing costs. We are able to solve a model that allows quite general technology for building houses and for preferences and which endogenises regional population density, land and house prices and commuting patterns in general equilibrium. We find that so long as improvements in travel technology proceed at a pace that is in a fixed proportion (of one-half) to the growth in productive potential (the sum of labour force growth and general productivity growth) there is a BGP with no change in real house prices. But once travel improvements fall back we are no longer on a BGP and real house prices and rents rise. How fast they then rise becomes dependent on a range of parameters that play little role when travel improvements were at a rate near half GDP growth. We then find that plausible parameter estimates plugged into this growth model can easily generate ever rising housing costs—relative to the price of other goods *and* to incomes. But there is great sensitivity of that to parameters that reflect both preferences (between different characteristics of houses) and technology. One key technology factor is how one combines structures and land to create housing. That has changed—the New York skyline shows that it is now possible to erect super-tall residential buildings on small plots of land and squeeze more residential space from a plot than was possible in the past. Whether that can drive our parameter ε to higher levels that fundamentally change the likely future cost of housing is an interesting question.

Price sensitivity of demand for housing—reflecting the substitutability of housing for other goods and services in creating satisfaction—is another factor with a very powerful effect on the longer-term path of housing costs.

The tractable model can throw light on policy issues that arise from the rising cost of housing. One is the relative effectiveness in holding down the cost of housing, of improving travel infrastructure, and raising the substitutability between land and structures in creating housing. Our results suggest that increasing the land-structure substitutability, either in house building or in consumer preferences, is particularly effective in keeping house prices down while travel improvements may be less powerful than in the past.

Appendix A. Transport Speeds over Time

Detailed data on travel speeds is rather patchy, especially for the late nineteenth and early twentieth century. We rely on surveys from various industrial economies to piece together a plausible path for the overall evolution of travel speeds and use this to tie down the rate of change of λ.

Survey evidence suggests that in the UK in the late nineteenth century around 60% of journeys people undertook were made on foot (Pooley and Turnbull, 1999, table 5). Fifty years earlier—and before the spread of railways—the proportion walking would have been much higher. By 1920 a substantial proportion (around 50%) used trains or buses for journeys and the proportion walking had fallen to under 40%. For those travelling to London trains and buses accounted for close to 80% of journeys over 1920 to 1929 (Pooley and Turnbull, 1999, table 6). Heblich *et al.* (2018) report a slightly more than threefold increase in travel speeds in the UK between the early nineteenth century and 1920. Leunig (2006) reports an enormous (sixteenfold) increase in distances travelled in the UK between 1865 and 1912 and very large improvements in average travel speeds. The rise in travel speeds was particularly fast between 1860 and 1900. Green (1988) showed that travel distance to a fixed location of work in London more than doubled over this period. In the next 70 years, according to the data reported in Pooley and Turnbull (1999), travel speeds rose by about 80%; over this period car use rises greatly but the improvement in average travel speeds is very substantially smaller than when railways and buses displaced walking in the earlier period. For those travelling into London, which is clearly the economic centre of Britain, improvements in average travel speeds between 1925 and 2000 was markedly slower than the average for all journeys in the UK.

Based on this UK historical data the annual improvement in speeds over the period 1850–1925 was a little under 1.5% a year while over the period 1925–2000 it dropped to about 0.5% a year.

The pattern suggested by the UK data is consistent with what happened in the USA. Ausubel *et al.* (1998) shows something similar for the USA where walking and travel by horse account for a sharply falling share of travel over time, replaced first by trains and later by cars. They show estimates of the average daily distance travelled by US citizens over the period 1880 to 1998. If time spent travelling is roughly constant this is a measure of the evolution of average travel speed. The average annual rise over the whole period is 2.7%, but average travel times have increased so 2.7% is an overestimate of the rise in passenger travel speeds. The data also show a clear falling off in the growth of travel distances from about 1970 (Ausubel *et al.*, 1998, figure 3). The growth in distance is at its fastest between around 1900 and 1930. The common pattern of the spread of railways is one factor behind a general and sharp rise in speeds in the later part of the nineteenth century. Railroad density (kilometres of railways per 1,000 people) show enormous expansion in Britain and the USA between 1850 and 1920, when railroad density peaked, Hugill (1995). More recently, and now that car journeys have become the most used means of travel in developed countries, improvements in travel speeds have fallen. Evidence of flat or even rising travel times into large cities exists for many countries: Glaeser and Kohlhase (2004) present evidence of rising congestion and of travel times into major US cities since 1990; van Wee *et al.* (2006) present evidence of increasing travel times in Holland. Ausubel *et al.* (1998) presents some evidence that average car speeds in the USA have not changed since the World War II. In considering the future evolution of house prices it may be plausible to consider that λ is now constant.

Sources for estimates of travel speeds in England shown in Figure 1: We construct the measure of the average passenger speed in England by taking estimates of the average speed each year by different means of transport and weight them by estimates of what proportion of journey are completed by each transport mode.

The modes of transport are: walking; horse (including carriages and horse omnibuses); rail (overground and underground); road vehicles.

For recent years (from 1988) shares of each mode of transport are from 'Commuting in England 1988–2015', UK Department of Transport. Snapshots of the use of different modes of transport for specific years in earlier periods are drawn from several sources. These data are interpolated to provide a continuous series for each transport mode. Data are drawn from: Gruebler (1990); Pooley and Turnbull (1999); Leunig (2006); Heblich *et al.* (2018); Pooley *et al.* (2018) and from the UK Government report 'The History of Transport Systems in the UK' (2018, Government Office for Science).

Data for speeds are drawn from largely the same sources. Leunig (2006) provides precise estimates of average railway travel speeds in the nineteenth and early twentieth centuries. Walking speeds are assumed constant at 2.5mph. Several sources quote horse and horse-drawn carriage speeds in the mid-to-late nineteenth century into cities as being not much above a fast walking pace. We set this at 4.5 mph. Car speeds have slowed in recent decades from levels that were around 20 mph in the 1960s. They increased steadily towards that level from when road transport first became significant at the start of the twentieth century.

Various cross checks on the plausibility of the overall estimates are applied. The figures generated for average speed since the 1980s match the data on average journey times and average distance travelled in the UK Government report 'Commuting in England 1988–2015'. Relative use of different means of transport in the nineteenth and early twentieth centuries (particularly trains) is consistent with data in Gruebler (1990) on infrastructure investment and capital stocks in different modes of transport and on his estimates of when horse-drawn transport essentially became unimportant (around 1930) and when road travel overtook trains (just after the World War II). The path of travel speeds in the late nineteenth century and up to the mid-twentieth century show rapid rises which match the data from England on average travel times (which have risen slightly) and average distances travelled (which increase greatly).

Appendix B. Proof of Existence of Balanced Growth Path

We first show that in a country of infinite size, ($l_{\max} = \infty$), there exists a balanced steady state. We then appeal to a Turnpike theorem to argue that when a country is of finite size, the economy will hug this optimal path until the urban expansion approaches the edge of the country. We proceed by assuming the economy is on BGP at t_0 and verify that if the travel tax falls at rate $(g+m)/2$ then economy remains on the BGP. Assume for all time $t \geq t_0$, supporting prices are constant; $r_t = r_{t_0}$, $w_t = w_{t_0}$ and $p_{0t} = p_{0t_0}$ where the interest rate r_t satisfies the BGP (or Ramsey) condition

$$r_t = r_{t_0} = \theta + \gamma\,(g+m).$$

To describe the spatial economy along the BGP, introduce the scaled location variable, $\widetilde{l}\,(l,t) = le^{-(g+m)(t-t_0)/2}$ which we denote as $\widetilde{l}_t$ for short. Along the BGP, this scaling maps the growing urban area back onto the urban area at t_0. This enables us to describe all variables in the spatial economy for $t \geq t_0$ in terms of their values at t_0. First, given constant rental prices at the CBD, $p^S_{0t} = p^S_{0t_0}$, and travel tax $\lambda_t = \lambda_{t_0} e^{-(g+m)(t-t_0)/2}$, then equation (4) implies that $p^S_{lt} = p^S_{\widetilde{l}_t t_0}$ at all locations; that is that the rental price at time t and location l is equal to the rental price at time t_0 and location $\widetilde{l}_t$. Similarly for the edge of the urban area, equations (12) or (13) imply that $l_{t_0,\text{Edge}} = \widetilde{l}\left(l_{t,\text{Edge}}, t\right)$; thus the urban area at time t maps back onto the urban area at time t_0. As distances are scaled back by $e^{-(g+m)(t-t_0)/2}$, areas will be scaled back by the square of this; hence we can write $R_{\widetilde{l}_t t_0} = R_{lt} e^{-(g+m)(t-t_0)}$. Under these definitions we show that the equilibrium conditions are satisfied at t if they are satisfied at t_0.

Equilibrium condition (23) says that demand for land equals the supply of land. If this is satisfied at t_0, then at time t we can map all quantities onto their respective values at t_0 to show that

$$\int_0^{l_{t,\text{Edge}}} \left(\frac{2\pi l}{R_{i(l),t}}\right) dl = \int_0^{l_{t_0,\text{Edge}}} \left(\frac{2\pi \widetilde{l}_t}{R_{i(\widetilde{l}_t),t_0}}\right) d\widetilde{l}_t = 1,$$

and so (23) is satisfied at time t. Similarly if $C_{i(l),t} = C_{i(\widetilde{l}_t),t_0} e^{(g+m)(t-t_0)}$, $S_{lt} = S_{\widetilde{l}_t t_0} e^{(g+m)(t-t_0)}$ and $B_{lt} = B_{\widetilde{l}_t t_0} e^{(g+m)(t-t_0)}$ then equations (17), (16) and (18) imply that $C_t = C_{t_0} e^{(g+m)(t-t_0)}$, $S_t = S_{t_0} e^{(g+m)(t-t_0)}$ and $B_t = B_{t_0} e^{(g+m)(t-t_0)}$.

Given the spatial economy is consistent with a BGP, it is trivial to show the rest of the economy is too. As both total consumption and housing grow at the rate $(g + m)$ then equations (20) and (22) are satisfied at t as they are satisfied at t_0. Finally as aggregate capital K is also growing at rate $g + m$, then the production constraint (6) as well as the budget constraint (5) are all satisfied at t as they are satisfied at t_0.

Thus in a country of infinite size, there exists a steady state BGP. For a country of finite size, it will be optimal to converge towards this BGP as close as possible for as long as possible before the constraint of the fixed factor (land) forces the economy to significantly diverge from this path. Our numerical simulations have this property when λ_t falls at the rate $(g + m)/2$.

Appendix C. Solution Technique

We solve a discrete time approximation to the continuous time model described in this paper using the relaxation approach first described in Laffargue (1990) and Boucekkine (1995). We let a period be a year, so that $t = 0, 1, 2, \ldots, T$ and solve for B_t and K_t at these points. To demonstrate that the path for these state variables describes the economy, note that given such a path one can calculate investment I_t^K and I_t^B from discrete time equivalents of equations (7) and (8) respectively and C_t from the production equation (6). Then given the three aggregate quantities, C_t, B_t and that the dynastic population is equal to 1, one can solve for the triplet (Q_t, r_t, p_{0t}^S) such that the housing sector in Subsection 1.4 is in equilibrium.

However we do also need a terminal condition for the state variables at $t = T + 1$ in order to calculate I_t^K and I_t^B at $t = T$. To choose these terminal values, we appeal to the Turnpike theorem of McKenzie (1976). We can assume any reasonable values for the state variables at $T+1$; for the optimal growth path between our initial condition and this terminal value will 'hug' the optimal growth path of the infinite horizon problem as closely as possible for as long as possible before deviating off to the given terminal value. We therefore assume that $K_{T+1} = B_{T+1} = 100$ for a very large T and only report the growth path for $t \ll T$.[17]

To solve for the state variables B_t and K_t at $t = 0, 1, 2, \ldots, T$, we are solving for $2(T + 1)$ unknowns and therefore need to $2(T + 1)$ constraints. The first constraint is the initial condition,

[17] In practise we solve the model for 450 annual periods, and report answers only for the first 250. We checked that the path over these first 250 periods was different by less than 10^{-5} if we solved the model instead over 600 periods.

that $B_0 + K_0$ equals initial stock of capital. A further T constraints stem from the dynamic efficiency condition, equation (20), at $t = 1, 2,\ldots,T$. The other $(T + 1)$ constraints are that the marginal product of residential buildings equals the marginal product of reproductive capital, equation (22) and (21) at $t = 0, 1, 2,\ldots,T$. Hence, along the optimal path, the model variables as described by that path of B_t and K_t must solve these $2(T + 1)$ constraints. We write these conditions as a set of non-linear equations denoted $f(B_t, K_t) = 0$ where f is a $2(T + 1)$-vector.

The relaxation approach starts from an initial guess for the path of the state variables.[18] It then uses a standard Newton–Raphson iterative procedure to solve the $2(T + 1)$ non-linear equations. We set the convergence condition to a change of less than 10^{-6} between iterations. To achieve this level of accuracy took no longer than a couple of minutes of a Intel Core i5 2.7GHz processor.

Imperial College, London and CEPR
Imperial College London

Additional Supporting Information may be found in the online version of this article:

Replication Package

References

Abelson, P. (1997). ‘House and land prices in Sydney from 1931 to 1989’, *Urban Studies*, vol. 34(9), pp. 1381–400.
Ahlfeldt, G.M. and McMillen, D.P. (2014). ‘New estimates of the elasticity of substitution of land for capital’, in *Regional Development and Globalisation: Best Practices*, 54th Congress of the European Regional Science Association, Louvain-la-Neuve.
Ahlfeldt, G.M. and Wendland, N. (2011). ‘Fifty years of urban accessibility: the impact of the urban railway network on the land gradient in Berlin 1890–1936’, *Regional Science and Urban Economics*, vol. 41(2), pp. 77–88.
Albouy, D., Ehrlich, G. and Liu, Y. (2014). ‘Housing demand and expenditures: how rising rent levels affect behaviour and cost-of-living over space and time’, in CEP/LSE Labour Seminar, vol. 28, London: University of Illinois.
Albouy, D., Ehrlich, G. and Liu, Y. (2016). ‘Housing demand, cost-of-living inequality, and the affordability crisis’, NBER Working Paper 22816, National Bureau of Economic Research.
Alonso, W. (1960). ‘A theory of the urban land market’, *Papers in Regional Science*, vol. 6(1), pp. 149–57.
Alonso, W. (1964). *Location and Land Use. Toward a General Theory of Land Rent*, Cambridge, MA: Harvard University Press.
Angel, S., Blei, A., Parent, J., Lamson-Hall, P. and Sïnchez, N. (2016). *Atlas of Urban Expansion*, vol. 1, Areas and Densities, New York: New York University, Nairobi: UN-Habitat, and Cambridge, MA: Lincoln Institute of Land Policy.
Atack, J. and Margo, R.A. (1998). ‘“Location, location, location!” The price gradient for vacant urban land: New York, 1835 to 1900’, *The Journal of Real Estate Finance and Economics*, vol. 16(2), pp. 151–72.
Ausubel, J.H., Marchetti, C. and Meyer, P.S. (1998). ‘Toward green mobility: the evolution of transport’, *European Review*, vol. 6(2), pp. 137–56.
Black, D. and Henderson, V. (1999). ‘A theory of urban growth’, *Journal of Political Economy*, vol. 107(2), pp. 252–84.
Bolt, J., Inklaar, R., de Jong, H. and van Zanden, J.L. (2018). ‘Rebasing “Maddison”: new income comparisons and the shape of long-run economic development’, Maddison Project Working Paper 10, University of Groningen.
Boucekkine, R. (1995). ‘An alternative methodology for solving nonlinear forward-looking models’, *Journal of Economic Dynamics and Control*, vol. 19(4), pp. 711–34.
Campbell, J.Y. and Cocco, J.F. (2003). ‘Household risk management and optimal mortgage choice’, *Quarterly Journal of Economics*, vol. 118(4), pp. 1449–94.

[18] This guess is not critical. However our initial guess was constructed by assuming $K_t/B_t = 3$ for all t. We then set K_t equal to

$$K_t = \left(K_0 - \left(K_0 - \frac{K_{T+1}}{e^{g(T+1)}}\right)\frac{t}{T+1}\right)e^{gt},$$

where K_0 was chosen so that the initial condition was satisified and K_{T+1} was equal to our chosen terminal value for K_t. This guess was good enough to ensure quick convergence.

Campbell, J.Y. and Cocco, J.F. (2007). ‘How do house prices affect consumption? Evidence from micro data’, *Journal of Monetary Economics*, vol. 54(3), pp. 591–621.
Davis, M.A., Fisher, J.D. and Whited, T.M. (2014). ‘Macroeconomic implications of agglomeration’, *Econometrica*, vol. 82(2), pp. 731–64.
Davis, M.A. and Heathcote, J. (2005). ‘Housing and the business cycle’, *International Economic Review*, vol. 46(3), pp. 751–84.
Deaton, A. and Laroque, G. (2001). ‘Housing, land prices, and growth’, *Journal of Economic Growth*, vol. 6(2), pp. 87–105.
Ermisch, J.F., Findlay, J. and Gibb, K. (1996). ‘The price elasticity of housing demand in Britain: issues of sample selection’, *Journal of Housing Economics*, vol. 5(1), pp. 64–86.
Favilukis, J., Ludvigson, S.C. and Van Nieuwerburgh, S. (2017). ‘The macroeconomic effects of housing wealth, housing finance, and limited risk sharing in general equilibrium’, *Journal of Political Economy*, vol. 125(1), pp. 140–223.
Fujita, M., Krugman, P.R. and Venables, A.J. (2001). *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge, MA: MIT Press.
Glaeser, E.L. and Kohlhase, J.E. (2004). ‘Cities, regions and the decline of transport costs’, *Papers in Regional Science*, vol. 83(1), pp. 197–228.
Green, D.R. (1988). ‘Distance to work in Victorian London: a case study of Henry Poole, bespoke tailors’, *Business History*, vol. 30(2), pp. 179–94.
Grossmann, V. and Steger, T.M. (2016). ‘Das house-kapital: a theory of wealth-to-income ratios’, CESifo Working Paper 5844, Center for Economic Studies and Ifo Institute.
Gruebler, A. (1990). *The Rise and Fall of Infrastructure: Dynamics of Evolution and Technological Change in Transport*, Heidelberg: Physica-Verlag.
Hall, R.E. (1988). ‘Intertemporal substitution in consumption’, *Journal of Political Economy*, vol. 96(2), pp. 339–57.
Heblich, S., Redding, S.J. and Sturm, D.M. (2018). ‘The making of the modern metropolis: evidence from London’, NBER Working Paper 25047, National Bureau of Economic Research.
Henderson, J.V. (2014). *Economic Theory and the Cities*, Salt Lake City, UT: Academic Press.
Hugill, P.J. (1995). *World Trade since 1431: Geography, Technology, and Capitalism*, Baltimore, MD: JHU Press.
Jaccard, I. (2011). ‘Asset pricing and housing supply in a production economy’, *BE Journal of Macroeconomics*, vol. 11(1), p. 33.
Jordà, Ò., Knoll, K., Kuvshinov, D., Schularick, M. and Taylor, A.M. (2019). ‘The rate of return on everything, 1870–2015’, *Quarterly Journal of Economics*, vol. 134(3), pp. 1225–98.
Khan, M.A. and Piazza, A. (2012). ‘Turnpike theory: a current perspective’, in (Palgrave Macmillan, ed.), *The New Palgrave Dictionary of Economics*, Palgrave Macmillan: London.
Kiyotaki, N., Michaelides, A. and Nikolov, K. (2011). ‘Winners and losers in housing markets’, *Journal of Money, Credit and Banking*, vol. 43(2–3), pp. 255–96.
Knoll, K., Schularick, M. and Steger, T. (2017). ‘No price like home: global house prices, 1870–2012’, *American Economic Review*, vol. 107(2), pp. 331–53.
Krugman, P. (1980). ‘Scale economies, product differentiation, and the pattern of trade’, *American Economic Review*, vol. 70(5), pp. 950–9.
Krugman, P. (1991). ‘Increasing returns and economic geography’, *Journal of Political Economy*, vol. 99(3), pp. 483–99.
Laffargue, J.P. (1990). ‘Résolution d’un modèle macroéconomique avec anticipations rationnelles’, *Annales d’Economie et de Statistique*, vol. 17, pp. 97–119.
Leunig, T. (2006). ‘Time is money: a re-assessment of the passenger social savings from Victorian British Railways’, *Journal of Economic History*, vol. 66(03), pp. 635–73.
Levinson, D.M. and Kumar, A. (1994). ‘The rational locator: why travel times have remained stable’, *Journal of the American Planning Association*, vol. 60(3), pp. 319–32.
Lucas, R.E. (2001). ‘Externalities and cities’, *Review of Economic Dynamics*, vol. 4(2), pp. 245–74.
McDonald, J.F. (1981). ‘Capital-land substitution in urban housing: a survey of empirical estimates’, *Journal of Urban Economics*, vol. 9(2), pp. 190–211.
McKenzie, L.W. (1976). ‘Turnpike theory’, *Econometrica*, vol. 44(5), pp. 841–65.
McMillen, D.P. (1996). ‘One hundred fifty years of land values in Chicago: a nonparametric approach’, *Journal of Urban Economics*, vol. 40(1), pp. 100–24.
Mills, E.S. (1967). ‘An aggregative model of resource allocation in a metropolitan area’, *American Economic Review*, vol. 57(2), pp. 197–210.
Mills, E.S. (1972). *Studies in the Structure of the Urban Economy*, Washington, DC: RFF Press.
Muth, R.F. (1969). *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*, Chicago: University of Chicago Press.
Muth, R.F. (1971). ‘The derived demand for urban residential land’, *Urban Studies*, vol. 8(3), pp. 243–54.
Ogaki, M. and Reinhart, C.M. (1998). ‘Measuring intertemporal substitution: the role of durable goods’, *Journal of Political Economy*, vol. 106(5), pp. 1078–98.
Piazzesi, M. and Schneider, M. (2016). ‘Housing and macroeconomics’, *Handbook of Macroeconomics*, vol. 2, pp. 1547–640.

Pooley, C.G., Jones, T., Jones, H., Spencer, B. and Chatterjee, K. (2018). 'Family formation and everyday travel in Britain since c. 1850', *The History of the Family*, vol. 23(2), pp. 290–306.

Pooley, C.G. and Turnbull, J. (1999). 'The journey to work: a century of change', *Area*, vol. 31(3), pp. 281–92.

Quigley, J.M. and Raphael, S. (2005). 'Regulation and the high cost of housing in California', *American Economic Review*, vol. 95(2), pp. 323–8.

Redding, S.J. (2013). 'Economic geography: a review of the theoretical and empirical literature', in *Palgrave Handbook of International Trade*, pp. 497–531, London: Palgrave Macmillan.

Rognlie, M. (2016). 'Deciphering the fall and rise in the net capital share: accumulation or scarcity?', *Brookings Papers on Economic Activity*, vol. 2015(1), pp. 1–69.

Rossi-Hansberg, E. and Wright, M.L. (2007). 'Urban structure and growth', *Review of Economic Studies*, vol. 74(2), pp. 597–624.

Schafer, A. (2000). 'Regularities in travel demand: an international perspective', *Journal of Transportation and Statistics*, vol. 3(3), pp. 1–31.

Schafer, A. and Victor, D.G. (2000). 'The future mobility of the world population', *Transportation Research Part A: Policy and Practice*, vol. 34(3), pp. 171–205.

Smith, F.H. (2003). 'Historical evidence on the monocentric urban model: a case study of Cleveland, 1915–1980', *Applied Economics Letters*, vol. 10(11), pp. 729–31.

Thorsnes, P. (1997). 'Consistent estimates of the elasticity of substitution between land and non-land inputs in the production of housing', *Journal of Urban Economics*, vol. 42(1), pp. 98–108.

Van Nieuwerburgh, S. and Weill, P.O. (2010). 'Why has house price dispersion gone up?', *Review of Economic Studies*, vol. 77(4), pp. 1567–606.

van Wee, B., Rietveld, P. and Meurs, H. (2006). 'Is average daily travel time expenditure constant? In search of explanations for an increase in average travel time', *Journal of Transport Geography*, vol. 14(2), pp. 109–22.

Vissing-Jørgensen, A. (2002). 'Limited asset market participation and the elasticity of intertemporal substitution', *Journal of Political Economy*, vol. 110(4), pp. 825–53.

von Thünen, J.H. and Hall, P.G. (1966). *Isolated State: An English Edition of Der Isolierte Staat*, New York, NY: Pergamon.

Zahavi, Y. and Talvitie, A. (1980). 'Regularities in travel time and money expenditures', *Transportation Research Record*, vol. 750, pp. 13–9.