20/11/2017

Business cycles in space

Tom Holden[1], *Jonathan Swarbrick*[2]

Abstract: *Many important shocks in the real world are correlated not only in time, but also across some notion of space. This may be physical space, or the space of product, firm or household types. As a result of this spatial correlation, aggregate volatility emerges naturally from idiosyncratic shocks. In this paper, we introduce a tractable framework that allows for such shocks without necessitating the discretisation of space, or a departure from perturbation approximation. As a lead example, we construct a dynamic, stochastic, general equilibrium model of economic geography (DSGEEG). This model features population movement, firm dynamics and semi-endogenous growth. Using it, we show how transitory, spatially located productivity shocks can lead to permanent movements in population, helping to explain the decline of the U.S. mid-west. As an additional theoretical contribution, we derive conditions for the existence of continuous-in-space shock processes on a range of spaces of economic interest.*

Keywords: *new economic geography, heterogeneous agents, continuous in space, generalized Ornstein-Uhlenbeck processes processes*

JEL Classification: *E3, F4, C31*

[1] Contact address: School of Economics, FBEL, University of Surrey, Guildford, GU2 7XH, England.
Contact email: thomas.holden@gmail.com Web-site: http://www.tholden.org/
The author would like to thank Paul Levine and Antonio Mele for helpful discussions. Financial support provided to the author by the ESRC and the EC is also greatly appreciated. Furthermore, the author gratefully acknowledges the use of the University of Surrey FASS cluster for the numerical computations performed in this paper.
The research leading to these results has received funding from The European Union's Horizon 2020 Framework Programme for Research and Innovation, under grant agreement "Modelling of the Research and Innovation Policies" (MONROE, grant no. 727114).
[2] Contact address: Canadian Economic Analysis, Bank of Canada, 234 Wellington Street, Ottawa, ON, K1A 0G9, Canada.
Contact email: jswarbrick@bankofcanada.ca Web site: http://www.jonathanswarbrick.co.uk/
The views expressed in this paper are those of the authors and should not be interpreted as reflecting the views of the Bank of Canada.

1. Introduction

Modern dynamic macroeconomic models are driven by a variety of shocks, including shocks to productivity, discounting and the disutility of labour supply. Each of these is likely to be spatially correlated. Good ideas diffuse in the local neighbourhood, leading to spatial correlation in productivity. A desire to bring forward consumption, or to reduce labour supply might be driven by the weather, which is highly spatially correlated. Shocks may also be correlated across other notions of space, such as the space of product, firm or household types, or, more generally, any space of economic agents in which "nearby" agents are expected to share similar properties and experience similar shocks. For example, one might expect firms producing similar products to experience correlated shocks to their returns from R&D. Since such correlation allows idiosyncratic shocks to lead to aggregate volatility, this also gives a partial answer to the question of the sources of aggregate fluctuations.

We make two key contributions in this paper. Firstly, we build a DSGE model incorporating the key features of the new economic geography literature, driven by a continuous in space productivity process. This allows us to see how spatially located productivity shocks might drive movements in population, helping to explain, for example, the decline of the U.S. mid-west. Secondly, we introduce a tractable framework that allows for shocks that are continuous in space, and hence correlated over it. We derive conditions for the existence of such processes for a range of spaces of economics interest.

Our model features the key features highlighted by the new economic geography literature (see e.g. Krugman (1998) or Redding (2013) for reviews). Firstly, firms in the model produce differentiated varieties, and consumers have a taste for variety. Thus, in regions with high population, since there will be greater firm entry, there will also be greater productivity, via the variety effect. We model firm entry here following Bilbiie, Ghironi, and Melitz (2012). Since high population is associated with high productivity, high population regions are attractive to further inward migration. Concentration of population also reduces transport costs, further increasing effective productivity, particularly since we allow for a role for intermediates in production, following Krugman and Venables (1995).

These agglomeration forces are counterbalanced by the populations' need for living space, and their need to consume agricultural goods. As population increases in a location, more of it must be allocated for living space, and thus there is less remaining for the production of the agricultural good. Consequently, more of the agricultural good must be imported into the location, pushing up its relative price. This increases the desirability of locations producing significant quantities of the agricultural good. Additionally, an increase in productivity pushes up wages, and thus makes entry relatively more expensive in high population locations.

While these dispersion forces are enough for the existence of a steady-state with uniform population, it will turn out that they are insufficient for the local stability of that steady-state. Small, temporary, exogenous changes in productivity in one location can drive a move to an asymmetric steady-state. Thus, cities are an endogenous outcome of our model in steady-state, as in new economic geography models such as Krugman (1991). However, due to the difficulty in working with asymmetric steady-states, we work with a version of the model in which there is an additional dispersive force: a preference for living in a location with moderate population. We calibrate the strength of this force so that the model is only just locally stable. Under this calibration, positive productivity shocks have extremely persistent effects on population. In essence, the location that gets "lucky" originally will have a permanent advantage.

Apart from the specifics of our model, we also contribute to the modelling of stochastic processes that are continuous in space. For example, if space is modelled as a circle, then it may be natural to draw shocks across space from a Gaussian periodic Ornstein-Uhlenbeck process (Pedersen 2002). These may be thought of as the continuous time (here: space) equivalent of an autoregressive process, conditional on their values being identical at the ends of the unit interval. As a result, a single shock process generates both idiosyncratic and aggregate volatility, leading to both substantial heterogeneity across agents, as well as aggregate movement. We go on to generalise this construction to arbitrary compact spaces equipped with Radon measures.

The existing literature contains many techniques for solving heterogeneous agent models in which shocks are i.i.d. across agents. These generally necessitate time-consuming global solution methods (see e.g. den Haan 2010). Of course, one way of taking a local approximation to a heterogeneous agent model is to solve the model with a finite collection of agents. However, getting reasonable accuracy along these lines requires a prohibitively high number of agents, given the $\frac{1}{\sqrt{N}}$ rate guaranteed by the central limit theorem. In our modelling framework however, due to the continuity and bounded variation of the driving stochastic processes, accurate solution does not require large state spaces. This makes taking a perturbation solution to the model much more tractable, which is the approach we pursue here.

Interestingly, Desmet and Rossi-Hansberg (2014) write that: "Incorporating a continuum of locations into a dynamic framework is a challenging task for two reasons: it increases the dimensionality of the problem by requiring agents to understand the distribution of economic activity over time and over space, and clearing goods and factor markets is complex because prices depend on trade and mobility patterns. These two difficulties typically make spatial dynamic models intractable, both analytically and numerically." That our approach enables us to solve rich dynamic spatial models without drastic simplification is a considerable advantage to our approach.

1.1. The puzzle of aggregate volatility

Spatial correlation provides one partial explanation for the emergence of aggregate volatility. The standard puzzle is as follows: suppose an economy comprises N firms, each of which receive an i.i.d. productivity draw with finite variance. Then by the Lindeberg–Lévy CLT, as $N \to \infty$, the standard deviation of aggregate productivity declines as $\frac{1}{\sqrt{N}}$. Modern economies have millions of firms, so a back of the envelope calculation suggests they ought to have a miniscule variance. Our solution to the puzzle is almost trivial, we just assume that firms receive correlated shocks, as indeed they do in the real world. While perhaps not particularly "deep", this story both captures the micro-level data on correlation, resolves some macro-level puzzles, and provides a technical approach for addressing important macroeconomic issues.

This paper's story of the source of aggregate variation is a complement to those of Gabaix (2011) and Acemoglu et al. (2012). Gabaix (2011) argues that aggregate fluctuations may be explained if firms receive a multiplicative productivity shock with magnitude unrelated to their size and, the distribution of firm sizes is fat-tailed, so the economy contains large firms. Gabaix shows that in the extreme case of a s^{-1} tail to the firm size distribution, aggregate volatility declines as $\frac{1}{\log N}$, meaning that the observed aggregate volatility need not be particularly surprising.

However, this explanation can at best be part of the story, as the law of large numbers applies just as well within a large firm (comprised of many workers, in many factories, producing many different products or components) as it does across firms. So, we ought to be as surprised that the variance of productivity does not wash out in large firms, as we are that it does not wash out in aggregate. Gabaix suggests that the units that make up a firm may themselves follow a power-law size distribution, but the justification for this is unclear: if firms or firm units are receiving shocks with a second moment, it is hard to see how a power-law size distribution could emerge in the first place, given such a distribution has infinite second moment. Our model gives one way of completing Gabaix's story: if the many products produced by large firms are all nonetheless close in product space, then they will be tightly correlated, and there will still be substantial variance at the firm level.

1.2. Other related literature

There have been several prior dynamic models of economic geography, though these have usually been non-stochastic, with discrete space and simplifications which remove any forward-looking component to economic decisions. Examples preserving some dynamic component to decisions include the model of Caselli and Coleman (2001) and Eckert & Peters (2017), who build OLG models to explain structural change featuring capital investment decisions. Models with a static or purely backward looking solution include Michaels, Rauch, & Redding (2012), and Nagy (2016). Example with continuous space include Desmet & Rossi-Hansberg (2014) and Desmet, Nagy, and Rossi-Hansberg (2015), who propose models which, although dynamic, have a solution that is backwards looking. Examples of dynamic economic geography models with a stochastic component include Duranton (2007), who presents a version of the Grossman & Helpman (1991) model with a fixed discrete set of cities, and Rossi-Hansberg and Wright (2007), who produce a model with many point cities to match the city size distribution. In both cases, restrictive assumptions are made to ensure tractability.

On the macro side, there have been a few simple non-stochastic models involving continuous space, but without allowing for substantial interactions between locations. These include Brito (2004) and Boucekkine, Camacho, & Zou (2009), who present Ramsey models with continuous space, as well as Quah (2002), who presents a version of the Lucas (1988) model on the surface of a sphere.

Relative to the aforementioned papers, we are almost unconstrained in our model building. We will have continuous space, with a distribution of population over the space that is changing over time in response to opportunities. We will have both many state variables, and many forward-looking variables, and we will need to impose transversality constraints in solving the model. We also allow for a rich shock structure, with both spatially located shocks, and aggregate shocks.

Our model will be partly driven by spatially correlated productivity shocks. There has been a large literature finding the evidence for such spatial correlation. Some recent papers include Glass, Kenjegalieva, and Paez-Farrell (2013), Cardamone (2014), Glass, Kenjegalieva, and Sickles (2016).

1.3. Outline

The structure of our paper is follows. In Section 2, we introduce spatially correlated shock processes. We also present a first simple model with spatially correlated shocks, and discuss the tools we have produced to assist with the simulation of such models. The theoretical results on existence of continuous stochastic processes are relegated to Appendix A, due to their technical nature. In Section 2.3, we describe our dynamic stochastic general equilibrium model of economic geography (DSGEEG). Section 4 presents theoretical and computational results from this model. Section 5 concludes.

2. Spatially correlated shock processes

2.1. Fixing ideas

In macroeconomics, it is usually helpful to work with continuums of types, as it ensures individuals have no impact on aggregates. Given this, in order for nearby types to receive correlated shock draws, it is sufficient that the drawn shock is continuous in type space.

As an example, to fix ideas, suppose that firms produce products of types indexed by [0,1]. We would expect firms producing similar products to receive similar productivity shocks. We might then suppose that if $a_{x,t}$ is the log-productivity of firm $x \in [0,1]$ at t, then:

$$a_{x,t} = (1-\rho)\mu + \rho a_{x,t-1} + \sigma\varepsilon_{x,t},$$

where $\varepsilon_{x,t}$ is a continuous function of x. By induction, it is then trivial to show that $a_{x,t}$ is a continuous function of x as well, so firm productivity is always spatially correlated. Later we will consider generalisations of this structure in which productivity today can depend on the lagged productivity of nearby firms as well.

2.2. Simple examples

If we wanted a discrete time stationary stochastic process, using a Gaussian AR(1) process would be the natural choice. Ornstein-Uhlenbeck processes are the continuous equivalent of Gaussian AR(1) processes, and are defined on $\mathbb{R}$. Using a draw from an Ornstein-Uhlenbeck process is one possibility when we want shocks on e.g. the type space [0,1]. These processes are characterised by Gaussian marginals, with an auto-covariance function of the form:

$$\operatorname{cov}(\varepsilon_x, \varepsilon_{\tilde{x}}) = \sigma^2 \exp\{-\zeta|x-\tilde{x}|\},$$

where σ^2 scales the variance, and $\zeta > 0$ controls the persistence. An example of a realisation of such a process is given in Figure 1. As $\zeta \to 0$, we get Brownian motion, and as $\zeta \to \infty$ we get "white-noise". It turns out that Ornstein-Uhlenbeck processes are the unique stationary, Gaussian, Markovian process on $\mathbb{R}$ (Doob 1942). The other processes we look at in Appendix A will not be Markovian, but whereas the Markovian assumption is natural in time, in space it does not have any particular intuitive appeal.

One downside to using [0,1] as the type space is that types at the end of the interval may end up with different properties. For example, if space is physical space, and some goods are produced at each location, then households at the end of the interval will have to pay higher transport costs. It is often convenient then to work with spaces which are invariant under translation, since this will ensure that all points are a priori the same. One way to do this is to work with circles, spheres or torii. (Recall that a torus is a "donut" shape. It may be thought of as a square in which when you move off the top edge, you reappear on the bottom edge, and when you move off the left edge, you reappear on the right.)

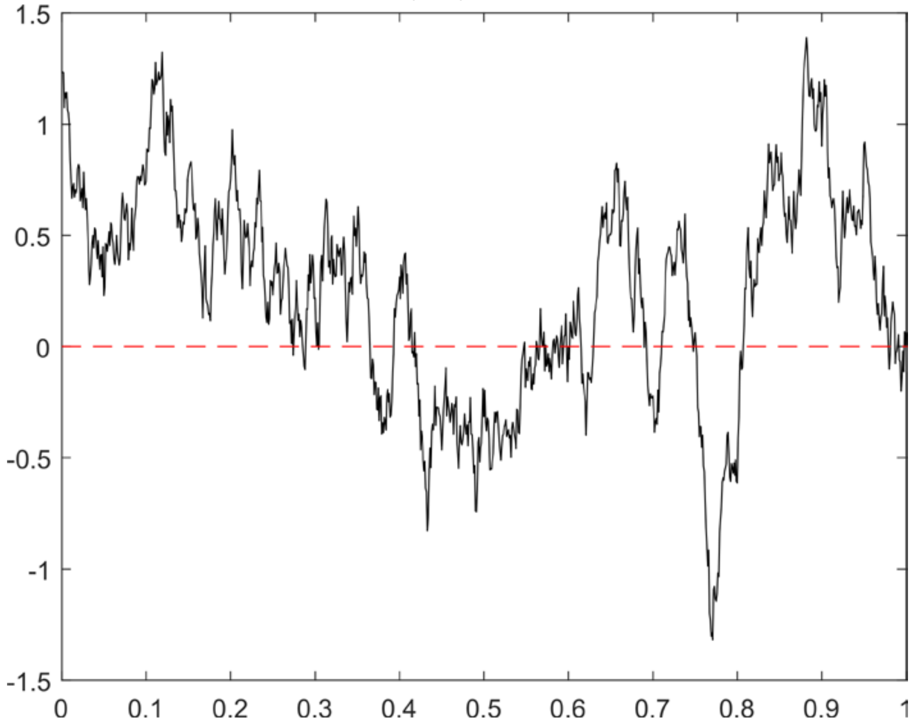

Figure 1: A realisation of an Ornsein-Uhlenbeck process with $\sigma = 1$ and $\zeta = 8$

It turns out that the natural Gaussian continuous stochastic process on a circle or a torus is characterised by an auto-covariance function of the form:

$$\operatorname{cov}(\varepsilon_x, \varepsilon_{\tilde{x}}) = \sigma^2 s\big(\zeta, d(x, \tilde{x})\big),$$

where σ^2 scales the variance, $\zeta > 0$ controls the persistence, d is the distance function (metric) being used on the circle (identified with [0,1]) or torus (identified with $[0,1] \times [0,1]$) and for all $\zeta, d > 0$:

$$s(\zeta, d) = \frac{\exp\big(-\zeta d + \zeta \bar{d}\big) + \exp\big(\zeta d - \zeta \bar{d}\big)}{\exp\big(\zeta \bar{d}\big) + \exp\big(-\zeta \bar{d}\big)},$$

where:

$$\bar{d} := \sup_{x, \tilde{x} \in X} d(x, \tilde{x})$$

is the maximum distance between points.

Further examples, along with proofs that these processes are well-defined are given in Appendix A.

2.3. A yeoman farmer model

To illustrate the basic idea of a DSGE model driven by a stochastic process that is continuous across space, we first present a very simple yeoman farmer model.

Suppose physical space is the circle, which we identify with [0,1]. At each point in space, there is a yeoman farmer, who owns the land at that point. We assume that there is no market in land, perhaps because owning land is necessary for survival. Production at a point $x \in [0,1]$ takes place with the production function $Y_{x,t} = A_{x,t}^{\alpha} L_{x,t}^{1-\alpha}$, where $A_{x,t}$ is the productivity of land at x and $L_{x,t}$ is the labour supply from the yeoman farmer at x. All farmers produce the same good, which is traded in a perfectly competitive market, and which we take as the numeraire. We suppose that there are no markets in state contingent securities, but that farmers can trade a one period zero net supply non-stochastic bond (an IOU).

Given this, farmers choose $L_{x,t}$ and bond holdings $B_{x,t}$ to maximise:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\log C_{x,t} - \frac{L_{x,t}^{1+\nu}}{1+\nu} \right]$$

subject to the constraint that:

$$C_{x,t} + B_{x,t} + \frac{\phi}{2} B_{x,t}^2 = Y_{x,t} + R_{t-1} B_{x,t-1},$$

where ϕ gives a cost of bond holdings, to ensure they are stationary. This leads to the first order conditions:

$$(1-\alpha) A_{x,t}^{\alpha} = L_{x,t}^{\alpha+\nu} C_{x,t},$$
$$1 + \phi B_{x,t} = \beta R_t \mathbb{E}_t \frac{C_{x,t}}{C_{x,t+1}}.$$

We may define aggregated variables by integration:

$$Y_t := \int_0^1 Y_{x,t}\, dx\,,\;\; C_t := \int_0^1 C_{x,t}\, dx\,,\;\; B_t := \int_0^1 B_{x,t}\, dx.$$

With these definitions, the market clearing condition is just that $B_t = 0$, which, providing $\phi \approx 0$, implies $Y_t \approx C_t$. Links between households in this model are quite weak, and entirely work via R_t.

It just remains for us to specify a stochastic process for $A_{x,t}$. In particular, we let $A_{x,t} = \exp a_{x,t}$, and suppose that:

$$a_{x,t} = \rho \left[(1-\chi) a_{x,t-1} + \chi \frac{\int_0^1 s(\eta, d(x,\tilde{x})) a_{\tilde{x},t-1}\, d\tilde{x}}{\int_0^1 s(\eta, d(0,\tilde{x}))\, d\tilde{x}} \right] + \sigma \varepsilon_{x,t},$$

where:

$$d(x,\tilde{x}) = \min\{|x-\tilde{x}|, 1-|x-\tilde{x}|\},$$

and where $\varepsilon_t \colon [0,1] \to \mathbb{R}$ is a draw from the mean zero Gaussian process with covariance:

$$\operatorname{cov}(\varepsilon_{x,t}, \varepsilon_{\tilde{x},t}) = s(\zeta, d(x,\tilde{x})),$$

where s is as defined in the previous section. This stochastic process features spatial spill-overs in productivity, controlled by χ and η. When $\chi = 0$, productivity at a location today is only a function of the shock and the previous value of productivity at that location. However, as χ increases, productivity becomes dependent on previous productivity at nearby locations, where η defines the meaning of "nearby".

Despite these spatial spill-overs, it turns out that aggregate log-productivity still follows an AR(1) process. To see this, define aggregate log-productivity by $a_t := \int_0^1 a_{x,t}\, dx$ and the aggregate shock by $\varepsilon_t := \int_0^1 \varepsilon_{x,t}\, dx$. Then by exchanging the order of integration, we have that:

$$a_t = \rho a_{t-1} + \sigma \varepsilon_t,$$

so average log-productivity follows a standard AR(1). It is also easy to see that $\varepsilon_t \sim N\left(0, \int_0^1 s(\zeta, d(0,\tilde{x}))\, d\tilde{x}\right)$, and that $\int_0^1 s(\zeta, d(0,\tilde{x}))\, d\tilde{x} < 1$, so the variance of aggregate productivity is lower than that of idiosyncratic productivity, in line with the data.

2.4. Simulating DSGE models with continuous in space stochastic processes

This model, like all the models we are interested in, possesses an infinite dimensional state, making simulation non-trivial. However, by the continuity of the shock, all variables including the states are continuous in space. Furthermore, it is easy to show that they are all of bounded variation. Consequently, their integrals may be approximated arbitrarily well by their values at finitely many points, via standard quadrature methods, and convergence of this quadrature will be much faster than with Monte Carlo (as used say in the Krusell-Smith algorithm). For example, if the trapezium rule is used on a circle, then the error will fall as $\mathrm{O}\left(\frac{1}{n^2}\right)$ (Rahman and Schmeisser 1990), rather than $\mathrm{O}\left(\frac{1}{\sqrt{n}}\right)$ with Monte Carlo.

These results carry across to convergence in distribution. For example, if we let $x \mapsto \varepsilon_{x,n}$ be the stochastic process on the circle in which $\left[\varepsilon_{\frac{a}{n},n}\right]_{a=1,\dots,n}$ is jointly normally distributed with mean zero and covariance $\operatorname{cov}\left(\varepsilon_{\frac{a}{n},n}, \varepsilon_{\frac{b}{n},n}\right) = \exp\left[-\zeta d\left(\frac{a}{n}, \frac{b}{n}\right)\right]$, and where $\varepsilon_{x,n}$ is given by linear interpolation away from these points, and we let $x \mapsto \varepsilon_x$ be the Gaussian stochastic process on the circle with mean zero and covariance $\operatorname{cov}(\varepsilon_{x,t}, \varepsilon_{\tilde{x},t}) = \exp[-\zeta d(x, \tilde{x})]$, then it is a theorem (Pedersen 2002) that $[\varepsilon_{x,n}]_{x\in[0,1]}$ converges in distribution to $[\varepsilon_x]_{x\in[0,1]}$ as $n \to \infty$, uniformly in x.

Our approach to simulation then is to choose a regular grid of points in space, and then to approximate the value of endogenous variables at points off this grid by linear interpolation. Having fixed the grid, we can then solve the model by standard methods for finite dimensional models; indeed, we may even use Dynare (Adjemian et al. 2011). Of course, manually adding equations for every point on the grid would be extremely time consuming. Luckily though, Dynare provides a pre-processor language that enables one to loop over points.

To assist further with the creation of spatial models, we provide a Dynare toolkit that can automatically define spatially correlated shock processes, including ones with spatial diffusion. It is available under an open source license from: https://github.com/tholden/DynareTransformationEngine. The model presented in this section is contained in "ExampleWithSpatialShocks.mod" in that repository. For a more complete example of all of the capabilities of this toolkit, see the code for this paper's main model here: https://github.com/tholden/DynamicSpatialModel.

2.5. Simulation results from our yeoman farmer model

We parameterise the model for quarterly periods as follows:

$$\alpha = 0.3, \beta = 0.99, \nu = 2, \rho = 0.95, \chi = 0.5, \eta = 8, \zeta = 4, \sigma = 0.02, \phi = 10^{-6}$$

and we use a 100-point approximation to the continuum,[3] and a second order perturbation approximation.

Figure 2 shows the sample paths from the final 1000 periods (250 years) of a 10,000 period (2,500 years) long simulation. Using the aforementioned Dynare toolkit, this simulation run took under 6 minutes. We graph both the aggregates and the cross-sectional distributions on the same plot. For the cross-sectional distribution, bright colours correspond to high-values. It is clear that there is substantial heterogeneity in consumption, chiefly driven by the accumulated bond holdings (which is close to following a random walk[4]), though shocks also generate short-term dispersion in consumption.

[3] As an indication of accuracy, note that with 50 points, the integral used in calculating $\int_0^1 s(\eta, d(0, \tilde{x}))\, d\tilde{x}$ is already accurate to 0.05%.

[4] This near random walk (with $\phi \approx 0$) is driven by the same factors that drive the standard random walk assets result in the open economy literature.

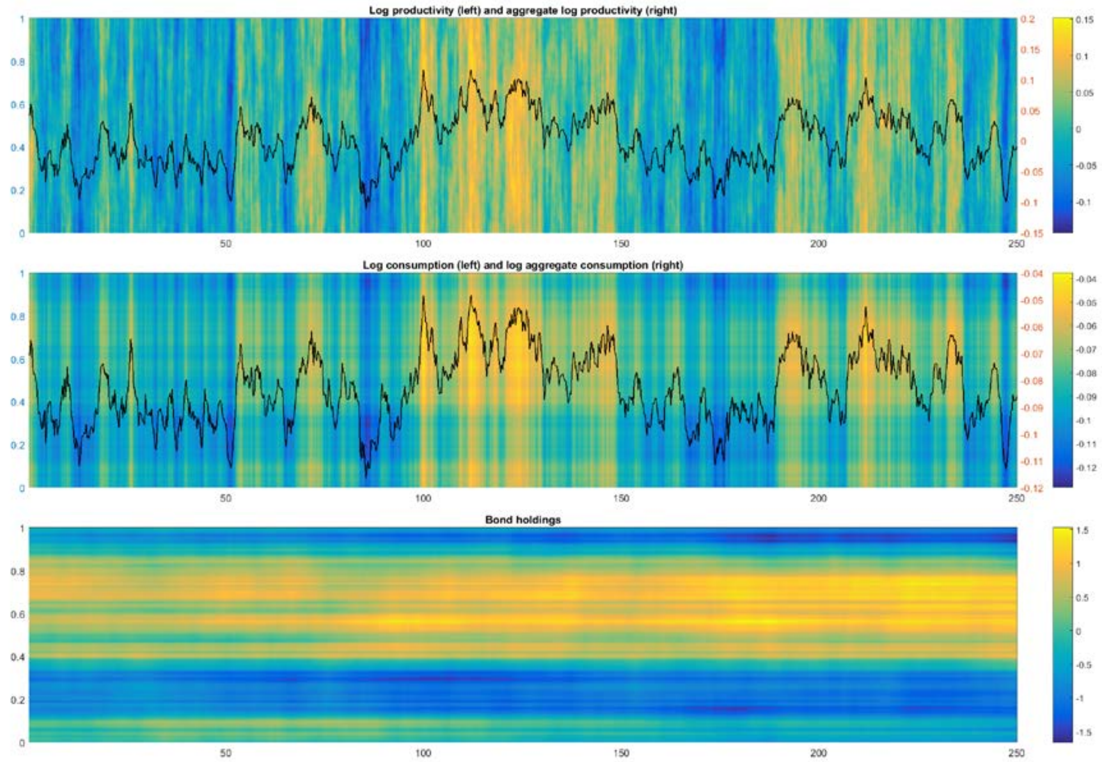

Figure 2: Simulation results from the yeoman farmer model, with $\chi = 0.5$.
The x-axis is measured in years. Bright colours are high values.

The effects of χ, which governs whether productivity responds to the lag of individual productivity ($\chi = 0$), or local productivity ($\chi = 1$), may be seen from Figure 3 and Figure 4. With $\chi = 0$, there is much greater spatial variance in productivity, which leads also to greater spatial variance in consumption.

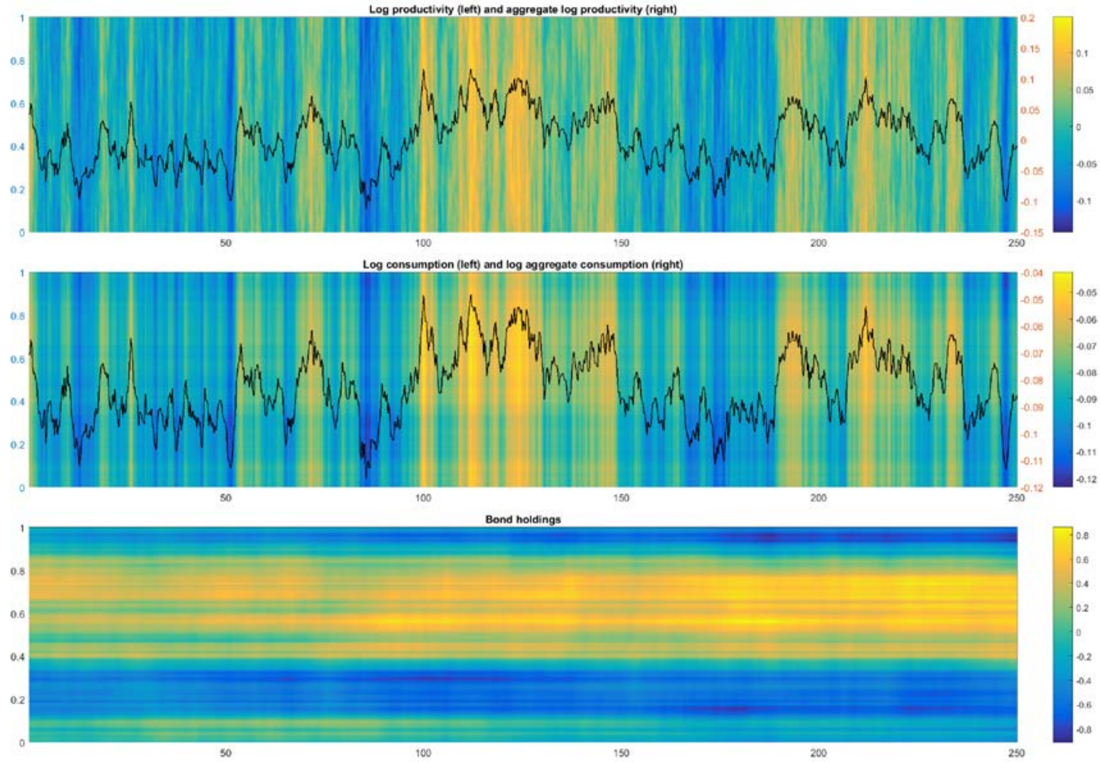

Figure 3: Simulation results from the yeoman farmer model, with $\chi = 1$.
The x-axis is measured in years. Bright colours are high values.

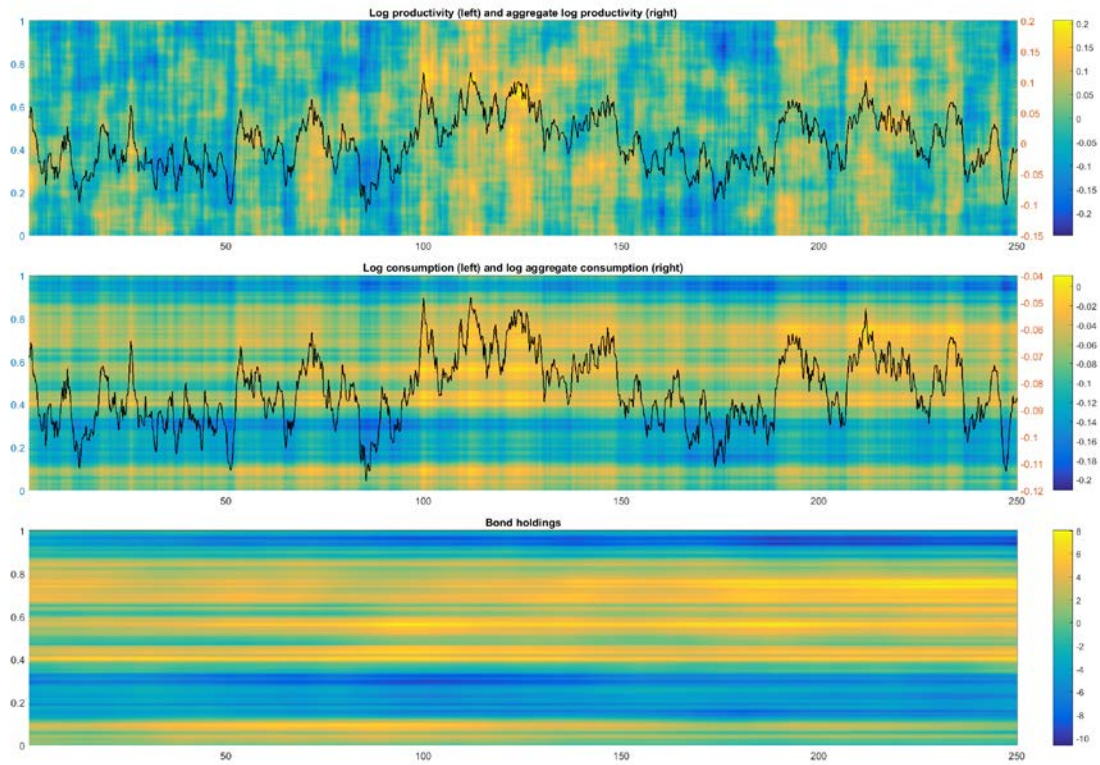

Figure 4: Simulation results from the yeoman farmer model, with $\chi = 0$.
The x-axis is measured in years. Bright colours are high values.

3. A dynamic stochastic general equilibrium economic geography model

We now present our DSGE model of economic geography. This model combines standard real business cycle features, with features from the workhorse models of new economic geography of e.g. Krugman (1991). The work of Bilbiie, Ghironi, and Melitz (2012) on firm dynamics is used to help bridge the gap between these two literatures. We are careful in our modelling choices to ensure that the model is consistent with balanced growth. This rules out non-homothetic preferences, for example.

The model features two types of final goods, agricultural products and manufactured products. Manufactured products are an aggregate of differentiated varieties produced by the firms in the model. Both these differentiated varieties and agriculture are produced using raw goods as an input, where raw goods are produced from capital, labour and intermediate inputs of manufactured goods. These raw goods may be thought of as providing production services. They are introduced chiefly to avoid complicating the model with multiple varieties of capital and labour. In agricultural production, the raw good is combined with land, whereas it is the sole input in the production of manufactured goods. For tractability, agricultural goods will be freely transportable and tradeable across locations, and manufactured goods will be untradeable. The differentiated varieties will be tradeable however, and will be subject to iceberg transportation costs.

Firms, capital and population will all have a density over space. We denote the set of points in space by X and assume that land is uniformly distributed over X. We normalise the total measure of X to 1, so $\int_X \mathrm{d}x = 1$. We assume that a metric is defined on X, giving the distance between any $x, \tilde{x} \in X$ as $d(x, \tilde{x})$.

At each location, there will be a representative household. For simplicity though, we assume that all these households are part of one representative family, and that household decisions are coordinated by a family head, who maximises a utilitarian social welfare function. As usual, this is equivalent to assuming the existence of complete markets between households. While assuming complete markets may be a little of a stretch, it greatly enhances the tractability of our model. Without this assumption, at each point in space there would be a distribution of asset holdings, as households who moved to that location would come with different assets to those who were already there. Furthermore, the decision of a household on where to move would be complicated by the need to consider what their utility would be at some location, which will differ in general from the utility of the households already there. If the reader is sceptical of the existence of complete markets in reality, it may help to think of our assumptions as giving the outcomes that a social planner could achieve with sufficient instruments. If real government policy is sufficiently close to optimal, then our model will provide a good guide to real world outcomes.

3.1. Manufactured good aggregator at $x \in X$

The non-tradeable manufactured final good at location x is produced by a perfectly competitive industry with access to the CES production function:

$$Y_{x,t} = \left[\int_X \int_0^{J_{\tilde{x},t}} \left(\frac{Y_{j,\tilde{x},x,t}}{\exp[\tau_t d(x,\tilde{x})]} \right)^{\frac{1}{1+\lambda}} dj \, d\tilde{x} \right]^{1+\lambda}. \tag{1}$$

Here, $\frac{1+\lambda}{\lambda}$ is the elasticity of substitution between varieties, τ_t gives the strength of iceberg transportation costs in period t, $J_{\tilde{x},t}$ gives the mass of firms located at $\tilde{x}$ in period t, and $Y_{j,\tilde{x},x,t}$ denotes the quantity of the differentiated variety produced by firm j at $\tilde{x}$ that is used in producing the final manufactured good at location x in period t. For convenience, we relabel firms each period, so that even with firm exit, the measure of firms located at $\tilde{x}$ in period t is still given by $[0, J_{\tilde{x},t}]$. Allowing for imperfect substitutability between varieties is crucial both because it allows for an increase in the measure of firms to increase productivity, providing an agglomerative force, and because it allows for the introduction of transport costs without having to deal with positivity constraints on consumption of varieties. Transport costs provide further agglomerative pressure, since in locations producing large numbers of varieties, less will need to be spent on transport.

In period t, the manufactured final good at x is sold at a price $P_{x,t}$, and the input produced by firm j in location $\tilde{x}$ is sold at a price $P_{j,\tilde{x},t}$. Thus, the profits of firms making the final manufactured good are given by:

$$P_{x,t}Y_{x,t} - \int_X \int_0^{J_{\tilde{x},t}} P_{j,\tilde{x},t}Y_{j,\tilde{x},x,t}\, dj\, d\tilde{x}\,.$$

From the first order condition for $Y_{j,\tilde{x},x,t}$ we then have that:

$$Y_{j,\tilde{x},x,t} = Y_{x,t}\left(\frac{P_{x,t}}{P_{j,\tilde{x},t}}\right)^{\frac{1+\lambda}{\lambda}} \exp\left[-\frac{\tau_t}{\lambda}d(x,\tilde{x})\right], \tag{2}$$

so demand is decreasing in the distance to the seller of the variety in question. From substituting equation (2) into equation (1), we have that:

$$P_{x,t} = \left[\int_X \int_0^{J_{\tilde{x},t}} \left(P_{j,\tilde{x},t}\exp[\tau_t d(x,\tilde{x})]\right)^{-\frac{1}{\lambda}} dj\, d\tilde{x}\right]^{-\lambda}. \tag{3}$$

Furthermore, equation (2) implies that the total demand for the good produced by firm j in location x at t is given by:

$$Y_{j,x,t} := \int_X Y_{j,x,\tilde{x},t}\, d\tilde{x} = P_{j,x,t}^{-\frac{1+\lambda}{\lambda}}\overline{Y}_{x,t}, \tag{4}$$

where:

$$\overline{Y}_{x,t} := \int_X Y_{\tilde{x},t}P_{\tilde{x},t}^{\frac{1+\lambda}{\lambda}} \exp\left[-\frac{\tau_t}{\lambda}d(\tilde{x},x)\right] d\tilde{x}.$$

3.2. Firms at $x \in X$

The j^{th} firm at location x producing a differentiated variety has access to the production function:

$$Y_{j,x,t} = Z_{j,x,t}, \tag{5}$$

where $Z_{j,x,t}$ is the amount of the raw good ("production services") it purchases in period t, at a price of $\mathcal{P}_{x,t}$. The firm maximises its profits which are given by:

$$Y_{j,x,t}(P_{j,x,t} - \mathcal{P}_{x,t}) = \left(P_{j,x,t}^{-\frac{1}{\lambda}} - \mathcal{P}_{x,t}P_{j,x,t}^{-\frac{1+\lambda}{\lambda}}\right)\overline{Y}_{x,t}.$$

From the first order condition for $P_{j,x,t}$, we derive the usual mark-up pricing condition:

$$P_{j,x,t} = (1+\lambda)\mathcal{P}_{x,t}. \tag{6}$$

Consequently, profits are equal across firms located at x in period t, and are given by:

$$\Pi_{x,t} := \frac{\lambda}{1+\lambda}(1+\lambda)^{-\frac{1}{\lambda}}\mathcal{P}_{x,t}^{-\frac{1}{\lambda}}\overline{Y}_{x,t}.$$

Furthermore, from substituting equation (6) into equation (3) we have that:

$$P_{x,t} = (1+\lambda)\left[\int_X J_{\tilde{x},t}(\mathcal{P}_{x,t}\exp[\tau_t d(x,\tilde{x})])^{-\frac{1}{\lambda}} d\tilde{x}\right]^{-\lambda}.$$

Much as in the model of Bilbiie, Ghironi, and Melitz (2012), firm entry requires paying ϕ_t units of the raw input, and firms exit at an exogenous rate, δ_J. Since firms are owned by the representative family, they discount the future with that family's, stochastic discount factor, which we denote by Ξ_{t+1}. This leads to the free entry condition:

$$\phi_t \mathcal{P}_{x,t} = \mathbb{E}_t \sum_{s=0}^{\infty} \left[\prod_{k=1}^{s} \Xi_{t+k} \right] (1-\delta_J)^s \Pi_{x,t+s},$$

i.e.:

$$\phi_t \mathcal{P}_{x,t} = \Pi_{x,t} + (1-\delta_J) \mathbb{E}_t \Xi_{t+1} \phi_{t+1} \mathcal{P}_{x,t+1}.$$

3.3. Capital holding company at $x \in X$

Without loss of generality, we assume that the capital stock at location x is owned by a representative capital holding company that is located there. The capital stock at x evolves according to:

$$K_{x,t} = (1-\delta_K) K_{x,t-1} + \left[1 - \Phi\left(\frac{I_{x,t}}{I_{x,t-1}} \right) \right] I_{x,t}, \tag{7}$$

where δ_K is the depreciation rate of capital, and Φ reflects Christiano, Eichenbaum, and Evans (2005) style investment adjustment costs, with $\Phi(1) = \Phi'(1) = 0$ and $\Phi''(1) > 0$. Capital is rented out at a rate $\mathcal{R}_{K,x,t}$ per unit at location x in period t and is immovable across locations. Including investment adjustment costs ensures that it is hard to move capital across locations by disinvesting in one location and reinvesting somewhere else. It thus helps to give persistence to the location of clusters of economic activity ("cities").

The representative capital holding company at x chooses period t investment to maximise their profits:

$$\mathbb{E}_t \sum_{s=0}^{\infty} \left[\prod_{k=1}^{s} \Xi_{t+k} \right] (\mathcal{R}_{K,x,t+s} K_{x,t+s-1} - P_{x,t} I_{x,t+s})$$

subject to law of motion for capital, equation (7). Writing $Q_{x,t}$ for the Lagrange multiplier on equation (7), this leads to the first order condition for $K_{x,t}$:[5]

$$1 = \mathbb{E}_t \Xi_{t+1} \frac{\mathcal{R}_{K,x,t+1} + Q_{x,t+1}(1-\delta_K)}{Q_{x,t}},$$

and first order condition for $I_{x,t}$:

$$P_{x,t} = Q_{x,t} \left(1 - \Phi\left(\frac{I_{x,t}}{I_{x,t-1}} \right) - \Phi'\left(\frac{I_{x,t}}{I_{x,t-1}} \right) \frac{I_{x,t}}{I_{x,t-1}} \right) + \mathbb{E}_t \Xi_{t+1} Q_{x,t+1} \Phi'\left(\frac{I_{x,t+1}}{I_{x,t}} \right) \left(\frac{I_{x,t+1}}{I_{x,t}} \right)^2.$$

3.4. Agriculture at $x \in X$

The agricultural sector at location x is perfectly competitive and has access to the production function:

$$F_{x,t} = L_{x,t}^{\gamma} Z_{F,x,t}^{1-\gamma},$$

where $Z_{F,x,t}$ is the amount of the raw good ("production services") used as an input to farming at location x in period t, and where $L_{x,t}$ is the amount of land allocated to farming at location x in period t. Farm land $L_{x,t}$ is rented from households at a rate of $\mathcal{R}_{L,x,t}$ per unit, and, as before, the raw good costs $\mathcal{P}_{x,t}$ per unit. We take the agricultural product as our numeraire (i.e. we assume it has unit cost), and further assume that it is tradeable without costs. The assumption of costless trade in agricultural products is common in the new economic geography literature previously cited. Introducing trade costs to the agricultural product would have required either introducing differentiation in agricultural products, or dealing with positivity constraints on agricultural production at each location, both of which would have substantially complicated the model. Furthermore, it is plausible that agricultural products should have relatively low trade costs

[5] The Lagrangian for this problem is contained in Appendix B.1.

compared to the rest of economic output, since many non-agricultural products are essentially non-tradeable (consider e.g. services, which, in our model, is subsumed within manufacturing).

Firms producing the agricultural good at location x in period t have profits:

$$F_{x,t} - \mathcal{R}_{L,x,t}L_{x,t} - \mathcal{P}_{x,t}Z_{F,x,t},$$

and thus choose $L_{x,t}$ such that:

$$\gamma \frac{F_{x,t}}{L_{x,t}} = \mathcal{R}_{L,x,t},$$

and $Z_{F,x,t}$ such that:

$$(1-\gamma)\frac{F_{x,t}}{Z_{F,x,t}} = \mathcal{P}_{x,t}.$$

3.5. Raw good production at $x \in X$

The raw good at location x is produced in period t by a perfectly competitive industry with access to the production function:

$$Z_{x,t} = \left[K_{x,t-1}^{\alpha}(A_{x,t}H_{x,t})^{1-\alpha}\right]^{1-\kappa}M_{x,t}^{\kappa},$$

where in period t capital $K_{x,t-1}$ is rented from capital holding companies at a rate of $\mathcal{R}_{K,x,t}$ per unit, labour $H_{x,t}$ is hired from the household at a wage $W_{x,t}$ per unit, and intermediate inputs of the final manufactured good, $M_{x,t}$, cost $P_{x,t}$ per unit. $A_{x,t}$ is productivity at location x in period t. Allowing for capital in production is important as high concentrations of capital are a defining feature of cities. It is also important to give a role for intermediate inputs of the final manufactured good in production, both because such inputs account for around half of gross output, and because this ensures that productivity is higher in locations where the final good is relatively cheap, generating further agglomerative pressure. We assume that the raw good is untradeable across locations, since it reflects production services.

Firms producing the raw good at location x in period t have profits:

$$\mathcal{P}_{x,t}Z_{x,t} - \mathcal{R}_{K,x,t}K_{x,t-1} - W_{x,t}H_{x,t} - P_{x,t}M_{x,t},$$

and thus choose $K_{x,t-1}$ such that:

$$(1-\kappa)\alpha\mathcal{P}_{x,t}\frac{Z_{x,t}}{K_{x,t-1}} = \mathcal{R}_{K,x,t},$$

$H_{x,t}$ such that:

$$(1-\kappa)(1-\alpha)\mathcal{P}_{x,t}\frac{Z_{x,t}}{H_{x,t}} = W_{x,t},$$

and $M_{x,t}$ such that:

$$\kappa\mathcal{P}_{x,t}\frac{Z_{x,t}}{M_{x,t}} = P_{x,t}.$$

3.6. Households and the representative family

There is a household with population $N_{x,t-1}$ at t at each $x \in X$. Population is pre-determined here to capture the fact that it takes time for people to move to exploit new opportunities elsewhere. As previously mentioned, for simplicity, we assume that all households are part of one representative family that takes decisions on their behalf.

In period t, the family head maximises the discounted utilitarian social welfare function:

$$\mathbb{E}_t \sum_{s=0}^{\infty}\left[\prod_{k=1}^{s}\beta_{t+k-1}\right]\int_X N_{x,t+s-1}\frac{U_{x,t+s}^{1-\varsigma}}{1-\varsigma}dx,$$

where β_t is the discount factor between periods t and $t+1$, $\varsigma \neq 1$[6] controls risk aversion and intertemporal substitution and:

$$U_{x,t} = \left(\frac{C_{x,t}}{N_{x,t-1}}\right)^{\theta_C} \left(\frac{E_{x,t}}{N_{x,t-1}}\right)^{\theta_F} \left(\frac{1-L_{x,t}}{N_{x,t-1}}\right)^{\theta_L} \left(\frac{1}{1+\nu}\Gamma^{1+\nu} - \frac{1}{1+\nu}\left(\frac{H_{x,t}}{N_{x,t-1}}\right)^{1+\nu}\right)^{\theta_H}$$
$$\cdot \left(\frac{1}{2}\Omega^2 - \frac{1}{2}\left(\log\left(\frac{N_{x,t-1}}{N_{t-1}}\right)\right)^2\right)^{\theta_N} \left(1-\frac{\mathcal{N}_{x,t}}{N_{x,t-1}}\right)^{\psi_1} \left(\bar{d} - \frac{\mathcal{D}_{x,t}}{\mathcal{N}_{x,t}}\right)^{\psi_2} \exp\left[\psi_3 \int_X \frac{N_{\tilde{x},t-1}}{N_{t-1}} \log \frac{\mathcal{N}_{x,\tilde{x},t}}{N_{x,t-1}} d\tilde{x}\right],$$

where:

$$N_{t-1} := \int_X N_{\tilde{x},t-1}\, d\tilde{x},$$
$$\mathcal{N}_{x,t} := \int_X \mathcal{N}_{x,\tilde{x},t}\, d\tilde{x},$$
$$\mathcal{D}_{x,t} := \int_X d(x,\tilde{x}) \mathcal{N}_{x,\tilde{x},t}\, d\tilde{x},$$

and:

$$1 = \theta_C + \theta_F + \theta_L + \theta_H + \theta_N + \psi_1 + \psi_2 + \psi_3. \tag{8}$$

The broad form of the utility function is dictated by the requirement that the model be consistent with balanced growth. This is particularly onerous in this model since the first order condition for population will include $U_{x,t}$, thus we cannot have additive terms within a household's felicity that have different growth rates. In order, the terms in $U_{x,t}$ are as follows:

- $\frac{C_{x,t}}{N_{x,t-1}}$ is consumption of the manufactured final good per head.
- $\frac{E_{x,t}}{N_{x,t-1}}$ is consumption ("eating") of the agricultural good per head.
- $\frac{1-L_{x,t}}{N_{x,t-1}}$ reflects the utility gained from access to unfarmed land. This captures the necessity of space for housing, which is otherwise unmodeled. While it might be natural to model a stock of housing, to keep the model tractable we abstract from the housing stock here.
- $\frac{1}{1+\nu}\Gamma^{1+\nu} - \frac{1}{1+\nu}\left(\frac{H_{x,t}}{N_{x,t-1}}\right)^{1+\nu}$ gives the utility gained from leisure. Here Γ controls the maximum amount of labour per head that will ever be supplied, and ν controls the (inverse) elasticity of labour supply.
- $\frac{1}{2}\Omega^2 - \frac{1}{2}\left(\log\left(\frac{N_{x,t-1}}{N_{t-1}}\right)\right)^2$ reflects the utility gained from living in a location with moderate population, with Ω controlling the maximum acceptable departure from this level. This term is necessary to ensure the stability of the symmetric steady-state.
- $1-\frac{\mathcal{N}_{x,t}}{N_{x,t-1}}$ reflects the utility of not having substantial outward migration $\mathcal{N}_{x,t}$, i.e. it captures people being upset by their friends and relatives moving away. This ensures that there is always an interior solution for $\mathcal{N}_{x,t}$, which is necessary for tractability.
- $\bar{d} - \frac{\mathcal{D}_{x,t}}{\mathcal{N}_{x,t}}$ gives the utility of not having migration to distant locations, where as in Section 2.2, $\bar{d} := \sup_{x,\tilde{x}\in X} d(x,\tilde{x})$ gives the maximum distance between points, and $\frac{\mathcal{D}_{x,t}}{\mathcal{N}_{x,t}}$ is a measure of the average distance moved. I.e. this term captures people being upset by their friends and relatives moving *far* away. This term helps to avoid rapid jumps in population to distant locations, implying that most migration will be between a city and its suburbs.

[6] The normal device of subtracting 1 from the numerator is not possible here, as it renders the first order condition for population inconsistent with balanced growth.

- $\exp\left[\int_X \frac{N_{\tilde{x},t-1}}{N_{t-1}} \log \frac{\mathcal{N}_{x,\tilde{x},t}}{N_{x,t-1}} d\tilde{x}\right]$ reflects a preference to have at least some migration to all locations ($\mathcal{N}_{x,\tilde{x},t}$ is the amount of migration from location x to location $\tilde{x}$ at t), with higher weight (and so higher migration) to locations with higher populations, i.e. it captures the inevitability of people starting new households with people from far away. This ensures that there is an interior solution for $\mathcal{N}_{x,\tilde{x},t}$, which is necessary for tractability.

The family head faces the budget constraint:

$$\int_X (P_{x,t}C_{x,t} + E_{x,t})\,dx + B_t = \int_X (\mathcal{R}_{L,x,t}L_{x,t} + W_{x,t}H_{x,t})\,dx + R_{t-1}B_{t-1} + \mathrm{T}_t,$$

where T_t includes all net profits from owning firms and capital holding companies, and where the family's bond holdings, B_t, are zero in equilibrium. The family head also faces the following constraint on the evolution of $N_{x,t}$, for all $x \in X$:

$$N_{x,t} = G_{N,t}N_{x,t-1} - \int_X \mathcal{N}_{x,\tilde{x},t}\,d\tilde{x} + \int_X \mathcal{N}_{\tilde{x},x,t}\,d\tilde{x}\,, \tag{9}$$

where $G_{N,t}$ is the growth rate of aggregate population N_t, $\int_X \mathcal{N}_{x,\tilde{x},t}\,d\tilde{x}$ is outwards migration from x and $\int_X \mathcal{N}_{\tilde{x},x,t}\,d\tilde{x}$ is inwards migration to x. Writing $\mu_{N,x,t}$ for the Lagrange multiplier on the law of motion for $N_{x,t}$, equation (6)(9), we may derive the following first order conditions[7] for consumption $C_{x,t}$:

$$\theta_C E_{x,t} = \theta_F P_{x,t} C_{x,t},$$

land $L_{x,t}$:

$$\theta_L E_{x,t} = \theta_F \mathcal{R}_{L,x,t}(1 - L_{x,t}),$$

hours $H_{x,t}$:

$$\theta_H \left(\frac{H_{x,t}}{N_{x,t-1}}\right)^{\nu} = \theta_F \frac{N_{x,t-1}}{E_{x,t}} W_{x,t} \left(\frac{1}{1+\nu}\Gamma^{1+\nu} - \frac{1}{1+\nu}\left(\frac{H_{x,t}}{N_{x,t-1}}\right)^{1+\nu}\right),$$

bonds B_t:

$$1 = \beta_t R_t \mathbb{E} \frac{N_{x,t}E_{x,t}U_{x,t+1}^{1-\varsigma}}{E_{x,t+1}N_{x,t-1}U_{x,t}^{1-\varsigma}},$$

population $N_{x,t}$:

$$\mu_{N,x,t} = \beta_t \mathbb{E}_t \Bigg[\mu_{N,x,t+1} G_{N,t+1} + U_{x,t+1}^{1-\varsigma} \\ + (1-\varsigma) U_{x,t+1}^{1-\varsigma} \Bigg[\theta_H \frac{\left(\frac{H_{x,t+1}}{N_{x,t}}\right)^{1+\nu}}{\frac{1}{1+\nu}\Gamma^{1+\nu} - \frac{1}{1+\nu}\left(\frac{H_{x,t+1}}{N_{x,t}}\right)^{1+\nu}} - \theta_N \frac{\log\left(\frac{N_{x,t}}{N_t}\right)}{\frac{1}{2}\Omega^2 - \frac{1}{2}\left(\log\left(\frac{N_{x,t}}{N_t}\right)\right)^2} + \psi_1 \frac{\mathcal{N}_{x,t+1}}{N_{x,t} - \mathcal{N}_{x,t+1}} \\ - (\theta_C + \theta_F + \theta_L + \psi_3) \Bigg] \Bigg],$$

[7] This is somewhat complicated by the need to differentiate with respect to functions. We solve this by first replacing expressions of the form $\int_X f(x)\,dx$ with $\frac{1}{|\tilde{X}|}\sum_{x\in\tilde{X}} f(x)$ where $\tilde{X} \subset X$ is a finite set. We then simplify, and take limits as $|\tilde{X}| \to \infty$ and as $\tilde{X}$ becomes dense in X. The Lagrangian, and further details on the derivation of these conditions is contained in Appendix B.1.

and migration $\mathcal{N}_{x,\tilde{x},t}$:

$$\mu_{N,x,t} = \mu_{N,\tilde{x},t} + (1-\varsigma)N_{x,t-1}U_{x,t}^{1-\varsigma}\left[\psi_3 \frac{N_{\tilde{x},t-1}}{N_{t-1}\mathcal{N}_{x,\tilde{x},t}} - \psi_1 \frac{1}{N_{x,t-1}-\mathcal{N}_{x,t}} - \psi_2 \frac{d(x,\tilde{x})\mathcal{N}_{x,t} - \mathcal{D}_{x,t}}{\bar{d}\mathcal{N}_{x,t}^2 - \mathcal{N}_{x,t}\mathcal{D}_{x,t}}\right].$$

We also have that the representative family's stochastic discount factor is given by:

$$\Xi_{t+1} := \beta_t \frac{N_{x,t}E_{x,t}U_{x,t+1}^{1-\varsigma}}{E_{x,t+1}N_{x,t-1}U_{x,t}^{1-\varsigma}},$$

and that $\frac{E_{x,t}}{N_{x,t-1}U_{x,t}^{1-\varsigma}} = \frac{E_{\tilde{x},t}}{N_{\tilde{x},t-1}U_{\tilde{x},t}^{1-\varsigma}}$ for all $x, \tilde{x} \in X$, implying that with $\varsigma > 1$, households with low utility have high food consumption (a pattern that certainly holds in the data). Note also that with $\varsigma > 1$, $\mu_{N,x,t}$ gives a measure of the undesirability of location x, so it will be optimal to reduce population in locations in which $\mu_{N,x,t}$ is high.

3.7. Market clearing

The final manufactured good is used for consumption $C_{x,t}$, investment $I_{x,t}$ and as an intermediate in raw good production $M_{x,t}$, giving the period t market clearing condition:

$$Y_{x,t} = C_{x,t} + I_{x,t} + M_{x,t}.$$

Since raw goods are used in farming, firm entry and by the producers of differentiated varieties, demand for raw goods in period t is:

$$\begin{aligned} Z_{x,t} &= Z_{F,x,t} + \phi_t\big[J_{x,t} - (1-\delta_J)J_{x,t-1}\big] + \int_0^{J_{x,t}} Z_{j,x,t}\,dj \\ &= Z_{F,x,t} + \phi_t\big[J_{x,t} - (1-\delta_J)J_{x,t-1}\big] + J_{x,t}(1+\lambda)^{-\frac{1+\lambda}{\lambda}}\mathcal{P}_{x,t}^{-\frac{1+\lambda}{\lambda}}\bar{Y}_{x,t}, \end{aligned}$$

where to derive the second line we have used equations (4) and (5). The agricultural product is only "eaten", and is freely traded across locations, giving the period t market clearing condition:

$$\int_X E_{x,t}\,dx = \int_X F_{x,t}\,dx.$$

3.8. Stochastic processes

We close the model by specifying the driving stochastic processes. We assume that productivity $A_{x,t}$ is driven by a permanent component that is not location specific, A_t^{P}, and a location specific transitory component, $A_{x,t}^{\mathrm{T}}$. In particular:

$$A_{x,t} = A_t^{\mathrm{P}} A_{x,t}^{\mathrm{T}},$$

where:

$$A_t^{\mathrm{P}} = G_{A,t}A_{t-1}^{\mathrm{P}},$$

and where $\log G_{A,t}$ follows the AR(1) process:

$$\log G_{A,t} = (1-\rho_{G_A})\log G_A + \rho_{G_A}\log G_{A,t-1} + \sigma_{G_A}\varepsilon_{G_A,t},$$

and $\log A_{x,t}^{\mathrm{T}}$ follows the spatial AR(1) process:

$$\log A_{x,t}^{\mathrm{T}} = \rho_{A^{\mathrm{T}}}\log A_{x,t-1}^{\mathrm{T}} + \sigma_{A^{\mathrm{T}}}\varepsilon_{A^{\mathrm{T}},x,t},$$

where $\varepsilon_{A^{\mathrm{T}},x,t}$ is a realisation of some continuous stochastic process on X. We also assume that the other aggregate stochastic variables follow AR(1) processes, with:

$$\begin{aligned} \log G_{N,t} &= (1-\rho_{G_N})\log G_N + \rho_{G_N}\log G_{N,t-1} + \sigma_{G_N}\varepsilon_{G_N,t}, \\ \log \tau_t &= (1-\rho_\tau)\log\tau + \rho_\tau\log\tau_{t-1} + \sigma_\tau\varepsilon_{\tau,t}, \\ \log\phi_t &= (1-\rho_\phi)\log\phi + \rho_\phi\log\phi_{t-1} + \sigma_\phi\varepsilon_{\phi,t}, \end{aligned}$$

and:

$$\operatorname{logit}\beta_t = (1-\rho_\beta)\operatorname{logit}\beta + \rho_\beta \operatorname{logit}\beta_{t-1} + \sigma_\beta \varepsilon_{\beta,t}.^{8}$$

4. Results

4.1. Growth rates

From combining the model's equilibrium conditions, it may be shown that the model admits a balanced growth path in which for any x, $Y_{x,t}$ has stochastic trend:

$$G_{\tilde{Y},t} \coloneqq (G_{A,t}G_{N,t})^{\frac{(1-\alpha)(1-\kappa)(1+\lambda)}{(1-\alpha)(1-\kappa)(1+\lambda)-\lambda}}.$$

(Note, this does not mean that $G_{\tilde{Y},t}$ will equal $\frac{Y_{x,t}}{Y_{x,t-1}}$ for any particular x. Rather, this means that $\frac{Y_{x,t}}{\tilde{Y}_t}$ will be stationary, where $\tilde{Y}_t$ evolves according to $\tilde{Y}_t = G_{\tilde{Y},t}\tilde{Y}_{t-1}$.) Since $(1-\alpha)(1-\kappa)(1+\lambda)-\lambda > 0$ in any reasonable calibration, this implies that the growth rate of output is higher than that of $G_{A,t}G_{N,t}$. Thus, this is a model of semi-endogenous growth. Exogenous growth in productivity or population leads to further endogenous growth since it increases the measure of firms producing differentiated varieties, which feeds into the love for variety embedded in our aggregator, equation (1). The smaller is λ, the weaker will be this endogenous growth channel, with purely exogenous growth in the $\lambda = 0$, perfect competition, limit. The presence of this channel is also suggestive of areas of high population ("cities") having higher productivity.

The stochastic trend of other variables may be given in terms of the stochastic trend in output. In particular, we have that for any x, the stochastic trend in $Z_{x,t}$ and $J_{x,t}$ is given by $G_{\tilde{Y},t}^{\frac{1}{1+\lambda}}$, the stochastic trend in $F_{x,t}$, $E_{x,t}$ and $\mathcal{R}_{L,x,t}$ is given by $G_{\tilde{Y},t}^{\frac{1-\gamma}{1+\lambda}}$, the stochastic trend in $W_{x,t}$ is given by $G_{\tilde{Y},t}^{\frac{1-\gamma}{1+\lambda}}G_{N,t}^{-1}$, the stochastic trend in $P_{x,t}$, $Q_{x,t}$ and $\mathcal{R}_{K,x,t}$ is given by $G_{\tilde{Y},t}^{-\frac{\gamma+\lambda}{1+\lambda}}$, the stochastic trend in $\mathcal{P}_{x,t}$ and $\Pi_{x,t}$ is given by $G_{\tilde{Y},t}^{-\frac{\gamma}{1+\lambda}}$, the stochastic trend in $\overline{Y}_{x,t}$ is given by $G_{\tilde{Y},t}^{-\frac{\gamma}{\lambda}}$, the stochastic trend in $U_{x,t}$ is given by $G_{\tilde{Y},t}^{\theta_C+\theta_F\frac{1-\gamma}{1+\lambda}}G_{N,t}^{-(\theta_C+\theta_F+\theta_L)}$, and the stochastic trend in $\mu_{N,x,t}$ is given by $G_{\tilde{Y},t}^{\left(\theta_C+\theta_F\frac{1-\gamma}{1+\lambda}\right)(1-\varsigma)}G_{N,t}^{-(\theta_C+\theta_F+\theta_L)(1-\varsigma)}$. Thus, amongst other results, the model predicts that the price of manufactured goods and capital is falling with respect to the price of agricultural goods, and that consumption of agricultural goods is growing less quickly than consumption of manufactured ones.

The multiple different stochastic trends in the model complicate its simulation. However, this is facilitated by the ability of the toolkit we provide here: : https://github.com/tholden/DynareTransformationEngine to automatically take care of detrending variables, once the stochastic trends are supplied.

4.2. Properties of the steady-state, and choice of space and spatial correlation

While the full steady-state of the detrended model does not admit a closed form solution, in the special case in which the space X is invariant under translation (i.e. X is a circle or a torus), then the detrended model admits a uniform steady-state in which all variables are constant over x, and in which some variables have a closed form solution. In particular, in the uniform solution in the absence of shocks, $L_{x,t} = \frac{\gamma\theta_F}{\theta_L+\gamma\theta_F}$ and $\frac{\mathcal{N}_{x,t}}{N_{t-1}} = \frac{\psi_3}{\psi_1+\psi_3}$. Thus, the steady-state amount of land used in agriculture is increasing in the importance of land for

[8] The logit function is defined by $\operatorname{logit} p = \log\left(\frac{p}{1-p}\right)$. Specifying the process in this way ensures that β_t remains in the unit interval.

agricultural production, and in the importance of food for utility, and decreasing in the importance of land for utility. Additionally, the steady-state amount of migration is increasing in the family's desire to have at least some migration to each location, and decreasing in the amount the family dislikes any migration.

When the space X is not invariant under translation, as in the case when X is the plane $[0,1] \times [0,1]$ with the usual Euclidean metric, numerical results suggest that the steady-state features a significant concentration of population around the centre, $\left(\frac{1}{2},\frac{1}{2}\right)$. To see why this is unsurprising, suppose that population were initially uniformly distributed. Then the centre would have lower average transport costs, since it is on average closer to other places. These lower transport costs would imply lower prices and higher productivity in the centre, making it an attractive destination for migration.

In reality, in the U.S. at least, we see a lot of population on the coasts, and less in the centre. This is partly down to historical artefact, as the coasts were settled first, and partly due to the fact that in reality the coasts have low transport costs both to other points on the coast, and to the rest of the world. Rather than modelling trade along the coast, and to the rest of the world, we keep things simple by modelling space as a torus, identified with $[0,1] \times [0,1]$. I.e. our model of space is similar to how the continental U.S. would be were it the case that if you crossed the Canadian border, you teleported onto an equivalent point on the Mexican border, and if you stepped off the pier in Boston, you teleported to Seattle. This also means that there is a uniform steady-state, which greatly eases its computation. We place the usual Euclidean norm on the torus, i.e.:

$$d([x_1,x_2],[\tilde{x}_1,\tilde{x}_2]) = \sqrt{(\min\{|x_1-\tilde{x}_1|,1-|x_1-\tilde{x}_1|\})^2 + (\min\{|x_2-\tilde{x}_2|,1-|x_2-\tilde{x}_2|\})^2}.$$

Given this, it makes sense to use the "natural" continuous stochastic process on the torus introduced in Section 2.2, so:

$$\operatorname{cov}\left(\varepsilon_{A^{\mathsf{T}},x,t},\varepsilon_{A^{\mathsf{T}},\tilde{x},t}\right) = s\left(\zeta,d(x,\tilde{x})\right),$$

where s is as defined in Section 2.2.

4.3. Calibration and parameterisation

We now describe the calibration and parameterisation that we use for our simulation results. For this exercise, and due to the cost of solving the model, we choose to calibrate only a few key parameters, setting others to values from the literature or reasonable values. U.S. evidence suggests that the average home buyer stays in their house for around 13 years.[9] Since this omits renters who likely move more often, we calibrate ψ_3 to hit a proportion of $\frac{1}{12.5\times 4} = \frac{1}{50}$ of household members moving each quarter. This implies that $\psi_3 = \psi_1 \frac{0.02}{1-0.02}$. U.S. evidence also suggests that the share of land with broadly agricultural usage is around 75%.[10] This requires us to set $\theta_L = \frac{1-0.75}{0.75}\gamma\theta_F$. Furthermore, in the U.S. food is around 20% of personal consumption expenditure excluding housing.[11] Thus we set $\theta_F = \frac{1}{4}\theta_C$. Finally, we note that population per km^2 for the contiguous US is 41.5, whereas for Wyoming (the least dense state) it is 2.33 for New Jersey (the most dense

[9] See http://eyeonhousing.org/2013/01/latest-study-shows-average-buyer-expected-to-stay-in-a-home-13-years/

[10] Data from https://www.ers.usda.gov/data-products/major-land-uses/, Summary Table 1. We classify cropland, grassland, pasture, range and forest-use land as agricultural, and the rest as non-agricultural.

[11] Data from https://www.bea.gov/iTable/iTable.cfm?reqid=19&step=2#reqid=19&step=3&isuri=1&1910=x&0=-9&1921=survey&1903=65&1904=2015&1905=2017&1906=a&1911=0.

state) it is 470.[12] These correspond to absolute log ratios to the whole U.S. of 2.88 and 2.43 respectively. We thus set $\Omega = 3 > 2.88$ so that such dispersion is not ruled out.

We further parameterise $\theta_H = \theta_C$, and $\psi_1 = \psi_2 = \frac{\theta_F}{2}$. With these restrictions, and the adding up constraint, equation (8), we just have one remaining degree of freedom for the utility share parameters, which pins down the relative size of θ_N. Numerical experiments reveal that when $\theta_N = 0$, the symmetric steady-state is locally unstable. This result is not too surprising given that the non-existence of symmetric steady-states is a common finding in new economic geography models, at least for some parameters. As a result, we choose θ_N high enough such that the model is stable, but low enough that were θ_N significantly smaller, the model would be unstable. The resulting model will be close to having a unit root, giving dynamics that illustrate the likely behaviour with θ_N smaller.

The final full set of parameters is as follows:

$$\alpha = 0.3, \gamma = 0.5, \kappa = 0.5, \nu = 2, \varsigma = 1.5, \zeta = 8, \lambda = 0.1, \delta_J = 0.01, \delta_K = 0.03, \Gamma = 1, \Omega = 3, \Phi''(0) = 4,$$

$$\theta_C = 0.2618, \theta_F = 0.0655, \theta_L = 0.0109, \theta_H = 0.2618, \theta_N = 0.3338, \psi_1 = 0.0327, \psi_2 = 0.0327, \psi_3 = 0.007,$$ [13]

$$G_A = 1.005, G_N = 1.0025, \tau = 1, \phi = 1, \beta = 0.99,$$
$$\rho_{A^\intercal} = 0.9, \rho_{G_A} = 0.8, \rho_{G_N} = 0.5, \rho_\tau = 0.95, \rho_\phi = 0.95, \rho_\beta = 0.95,$$
$$\sigma_{A^\intercal} = \sigma_{G_A} = \sigma_{G_N} = \sigma_\tau = \sigma_\phi = \sigma_\beta = 0.001.$$

4.4. Impulse responses

To understand the dynamic behaviour of our model, we start by simulating impulse responses. The code we used both to simulate impulse responses, and to simulate stochastic runs is available from: https://github.com/tholden/DynamicSpatialModel. This repository also includes the full set of these results, including videos showing the evolution over time of all distributions. In all of the simulations reported here we used a grid with effective size 9×9, with the bottom row of grid points always agreeing with the top row, and the right column always agreeing with the left column.

We start by looking at the effects of a 1% spatial productivity shock. (Given that $\sigma_{A^\intercal} = 0.001$, this is a magnitude 10 standard deviations shock.) Since space is invariant under translation, without loss of generality we may focus on a shock that is centred on the point $\left(\frac{1}{2}, \frac{1}{2}\right)$. As shocks are correlated across locations, we take the matrix square root of the covariance matrix to determine the impulse at each location. The impact of such a shock is shown in Figure 5. Note that where an aggregate IRF is shown, this gives the IRF to the integral of the variable over X. In the density plots, bright colours represent high values.

[12] Land area data from https://www.census.gov/geo/reference/state-area.html, 2015 population estimates from the United States Census Bureau.

[13] All to four decimal places. For the precise values used, consult https://github.com/tholden/DynamicSpatialModel.

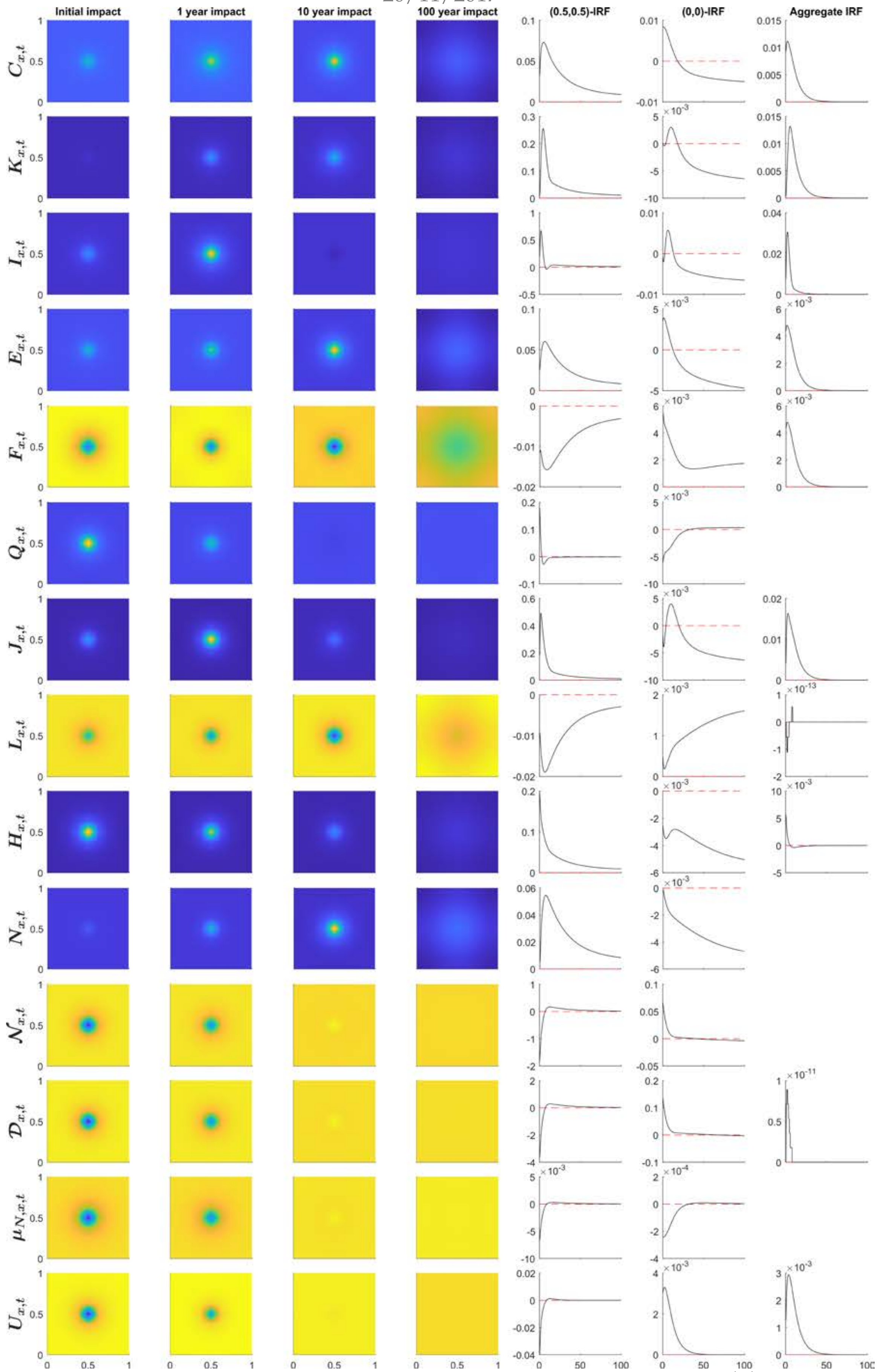

Figure 5: Impulse responses to a 1% spatial productivity shock centred on $(\frac{1}{2}, \frac{1}{2})$.

All y-axis values on IRFs are in percent. The x-axis is measured in years. Bright colours are high values.

The shock leads to an increase in consumption (manufactured and agricultural), investment, capital, the measure of firms, hours and population at the epicentre. Despite this, utility at the epicentre actually falls as people there are asked to work harder to take advantage of their high productivity. Since the increase in productivity leads to firm entry, it is optimal for the epicentre to move away from agricultural production, towards manufacturing. Consequently, agricultural production increases elsewhere, with a consequent increase in the land used for farming. This does not harm utility away from the epicentre since population flows towards the centre, reducing pressure on land in the periphery. As one would expect, aggregate utility increases overall from this positive productivity shock.

The effects of this initial shock are extremely persistent, with population still not back to trend 100 years after the initial shock. This suggests that our model can successful explain how small initial shocks can lead to city formation in one place, and not in another.

One explanation for the decline in the U.S. mid-west through the lens of the model is that the increase in productivity in e.g. San Francisco and New York has pulled people out of the mid-west and towards the coasts. That utility seems to have declined in the mid-west suggests that our model is still lacking important frictions, such as costs to adjust land usage.

In Figure 6 we plot 1% impulse responses to shocks to the models other driving processes. Note that all variables are shown relative to their stochastic trend, thus, for example, the short-run drop in consumption following a positive shock to $G_{A,t}$ is smaller in levels than appears from the picture, since the stochastic trend has permanently increased. Dynamics are broadly in line with dynamics from a modern business cycle model, suggesting that the main contribution of our model is in understanding what is going on within a country, rather than understanding aggregate dynamics per se.

4.5. Stochastic simulation

We finish by presenting selective results from a 1,000-year run of stochastic simulation from the model. Over this long period, most variables are fluctuating a great deal, and for many variables variance in time dominates variance across space. However, this is not the case for population. Figure 7 plots the distribution of population at 20-year intervals over this run, and a great deal of persistence can be seen over the entire span. Chance early shocks led to population accumulation in some regions, and these regions then become permanently more attractive, ensuring that those population changes persist. Thus, our model is capable of endogenously generating highly persistent clustering of population, i.e. cities. Nonetheless, there is still some movement in population. We see city centres drifting, cities, expanding, merging and joining, and we see a lot of movement between "cities" and their "suburbs".

To fully appreciate the dynamics of the model, we recommend watching the simulation videos for population and other variables that are contained within the model's repository: https://github.com/tholden/DynamicSpatialModel.

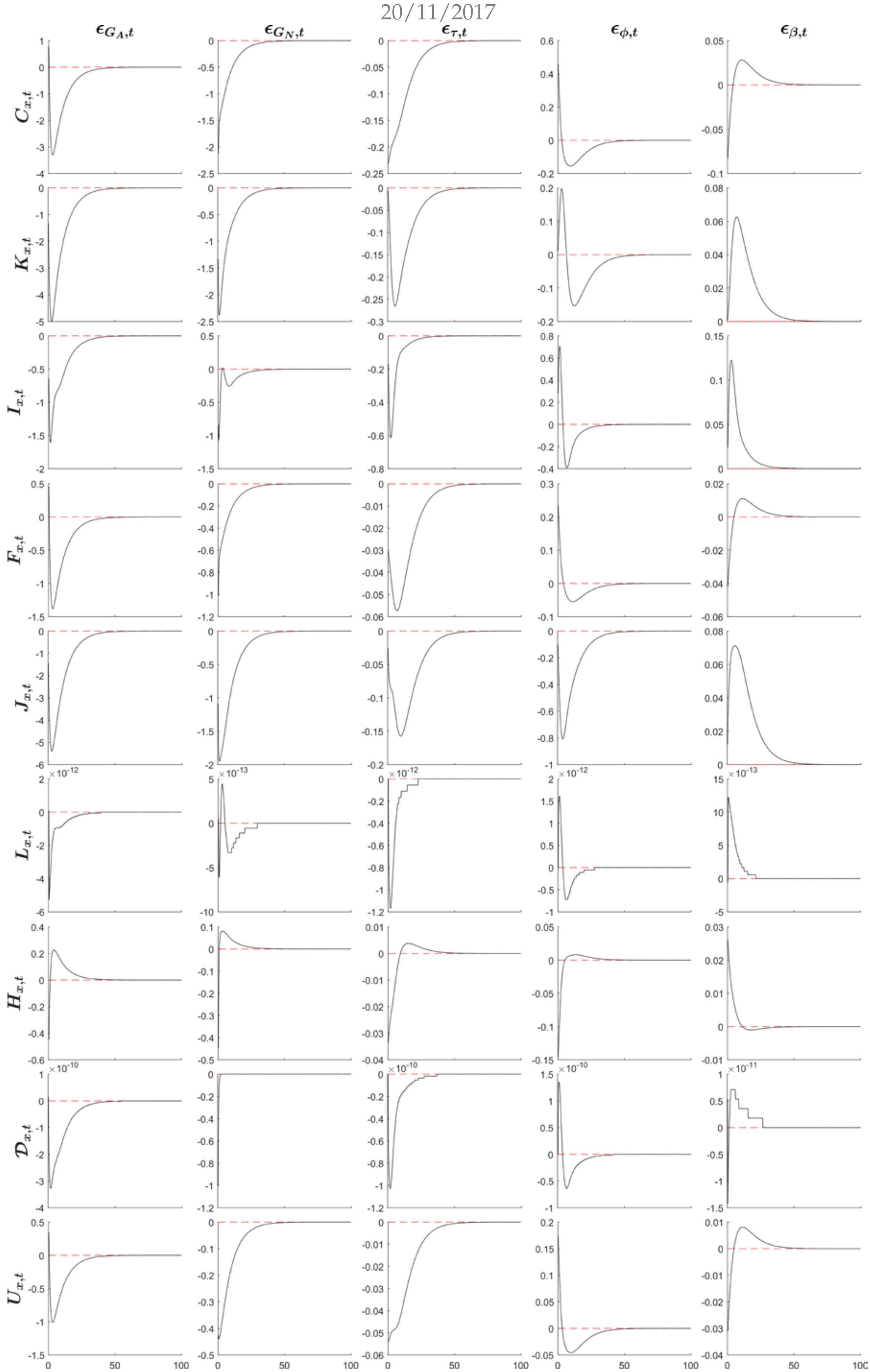

Figure 6: Aggregate impulse responses to 1% shocks to other stochastic processes

All variables are relative to their stochastic trends. All y-axis values are in percent. The x-axis is measured in years.

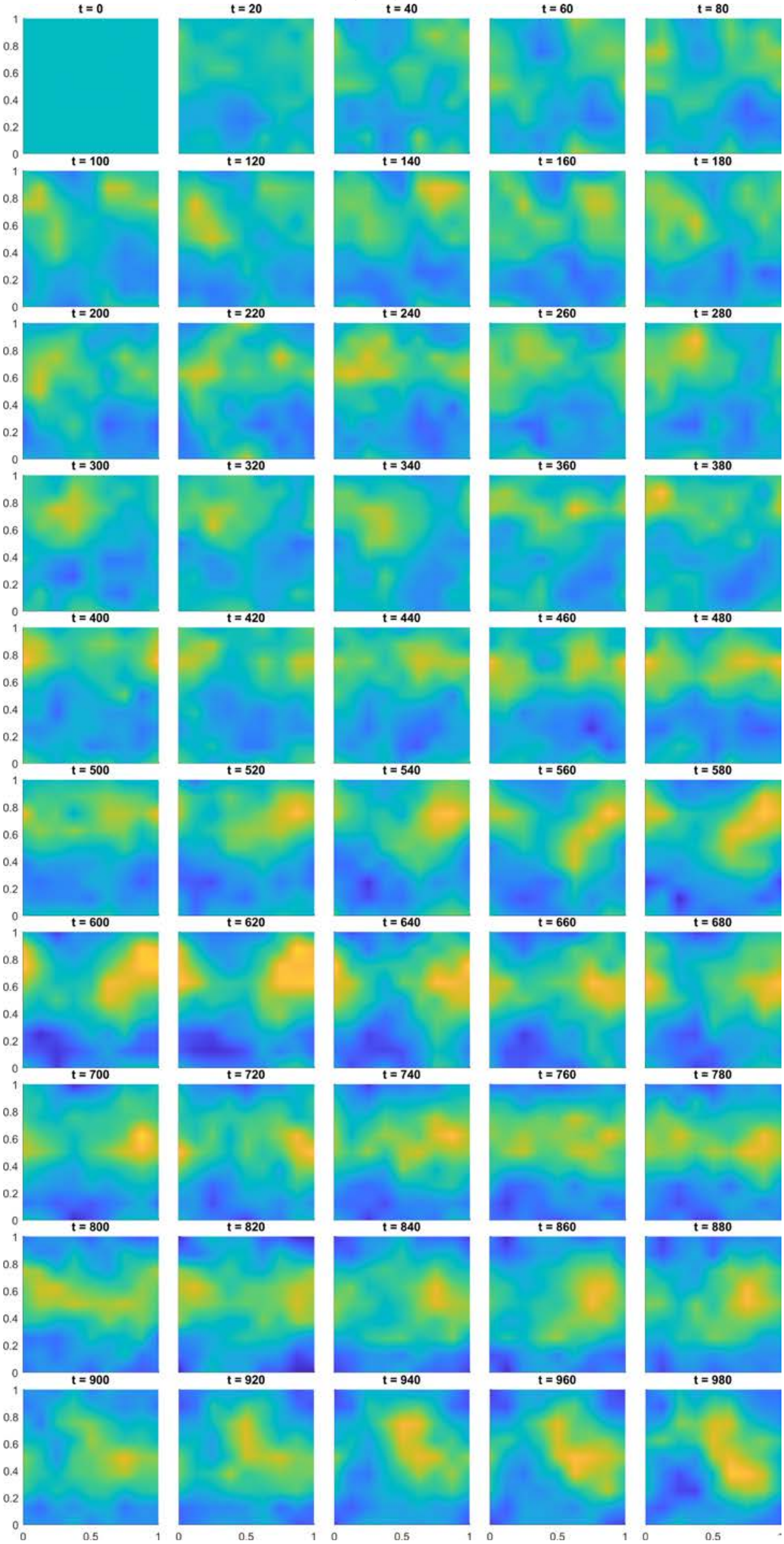

Figure 7: Population movements in a simulation run.
Time (above each plot) is measured in years.

5. Conclusion

This paper has presented a new approach to building heterogeneous agent macroeconomic models in which the heterogeneity is across space. While the paper focuses on applications in which space is physical space, our approach can also contribute to understanding heterogeneity across types, be it product variety, skills or preferences.

We suggested that spatial macroeconomic models should be driven by shocks that are continuous across space, and presented a variety of examples of such shock processes. We give further technical results on existence of such processes across a wide range of spaces of interest in Appendix A.

We went on to build a DSGE model featuring the key model components of the new economic geography literature. We showed that the model was able to generate extremely persistent movements in population, even given very strong preferences for a moderate population density. Thus, this is a model in which business cycle can shocks endogenously lead to the formation of new cities.

In future work, we plan on extending the model presented here, incorporating, for example, adjustment costs to land, that might ameliorate the need to have a preference for moderate population density. We will also undertake a more comprehensive calibration exercise, explore the asymmetric steady-states of the model, and assess the feasibility of solving the model at a higher order of approximation to capture the model's important non-linearities.

6. References

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. 2012. 'The Network Origins of Aggregate Fluctuations'. *Econometrica* 80 (5): 1977–2016.

Adjemian, Stéphane, Houtan Bastani, Fréderic Karamé, Michel Juillard, Junior Maih, Ferhat Mihoubi, George Perendia, Johannes Pfeifer, Marco Ratto, and Sébastien Villemot. 2011. *Dynare: Reference Manual Version 4*. CEPREMAP.

Bilbiie, Florin O., Fabio Ghironi, and Marc J. Melitz. 2012. 'Endogenous Entry, Product Variety, and Business Cycles'. *Journal of Political Economy* 120 (2) (April 1): 304–345.

Bogomolny, E, O Bohigas, and C Schmit. 2007. 'Distance Matrices and Isometric Embeddings'. *ArXiv Preprint ArXiv:0710.2063*.

Boucekkine, Raouf, Carmen Camacho, and Benteng Zou. 2009. 'Bridging The Gap Between Growth Theory And The New Economic Geography: The Spatial Ramsey Model'. *Macroeconomic Dynamics* 13 (01) (February): 20–45.

Brito, Paulo. 2004. *The Dynamics of Growth and Distribution in a Spatially Heterogeneous World*. Working Papers Department of Economics. ISEG - Lisbon School of Economics and Management, Department of Economics, Universidade de Lisboa.

Cardamone, Paola. 2014. *R&D, Spatial Proximity and Productivity at Firm Level: Evidence from Italy*. MPRA Paper. University Library of Munich, Germany.

Caselli, Francesco, and Wilbur John Coleman II. 2001. 'The U.S. Structural Transformation and Regional Convergence: A Reinterpretation'. *Journal of Political Economy* 109 (3) (June 1): 584–616.

Chepoi, Victor, Michel Deza, and Viatcheslav Grishukhin. 1997. 'Clin d'oeil on L 1-Embeddable Planar Graphs'. *Discrete Applied Mathematics* 80 (1): 3–19.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 2005. 'Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy'. *Journal of Political Economy* 113 (1): 1–45.

Desmet, Klaus, David Nagy, and Esteban Rossi-Hansberg. 2015. *The Geography of Development: Evaluating Migration Restrictions and Coastal Flooding*. CEPR Discussion Papers. C.E.P.R. Discussion Papers.

Desmet, Klaus, and Esteban Rossi-Hansberg. 2014. 'Spatial Development'. *American Economic Review* 104 (4): 1211–43.

Doob, J. L. 1942. 'The Brownian Movement and Stochastic Equations'. *Annals of Mathematics* 43 (2). Second Series (April 1): 351–369.

Duranton, Gilles. 2007. 'Urban Evolutions: The Fast, the Slow, and the Still'. *American Economic Review* 97 (1) (March): 197–221.

Eckert, Fabian, and Michael Peters. 2017. 'Spatial Structural Change and Agricultural Productivity'.

Gabaix, Xavier. 2011. 'The Granular Origins of Aggregate Fluctuations'. *Econometrica* 79 (3): 733–772.

Glass, Anthony J., Karligash Kenjegalieva, and Juan Paez-Farrell. 2013. 'Productivity Growth Decomposition Using a Spatial Autoregressive Frontier Model'. *Economics Letters* 119 (3): 291–295.

Glass, Anthony J., Karligash Kenjegalieva, and Robin C. Sickles. 2016. 'A Spatial Autoregressive Stochastic Frontier Model for Panel Data with Asymmetric Efficiency Spillovers'. *Journal of Econometrics* 190 (2): 289–300.

Gneiting, Tilmann. 2013. 'Strictly and Non-Strictly Positive Definite Functions on Spheres'. *Bernoulli* 19 (4): 1327–1349.

Gneiting, Tilmann, and Martin Schlather. 2004. 'Stochastic Models That Separate Fractal Dimension and the Hurst Effect'. *SIAM Review* 46 (2): 269–282.

Grossman, Gene M., and Elhanan Helpman. 1991. 'Quality Ladders in the Theory of Growth'. *The Review of Economic Studies* 58 (1) (January 1): 43–61.

den Haan, Wouter J. 2010. 'Comparison of Solutions to the Incomplete Markets Model with Aggregate Uncertainty'. *Journal of Economic Dynamics and Control* 34 (1): 4–27.

Hjorth, Poul, Petr Lisoněk, Steen Markvorsen, and Carsten Thomassen. 1998. 'Finite Metric Spaces of Strictly Negative Type'. *Linear Algebra and Its Applications* 270 (1): 255–273.

Koldobsky, Alexander. 1992. 'Schoenberg's Problem on Positive Definite Functions'. *ArXiv Preprint Math/9210207*.

Krugman, Paul. 1991. 'Increasing Returns and Economic Geography'. *Journal of Political Economy* 99 (3) (June): 483–499.

———. 1998. 'What's New about the New Economic Geography?' *Oxford Review of Economic Policy* 14 (2) (June 1): 7–17.

Krugman, Paul, and Anthony J. Venables. 1995. 'Globalization and the Inequality of Nations'. *The Quarterly Journal of Economics* 110 (4): 857–880.

Kuniewski, Sebastian P, and Jolanta K Misiewicz. 2014. 'Positive Definite Norm Dependent Matrices In Stochastic Modeling'. *Demonstratio Mathematica* 47 (1): 211–231.

Lucas, Robert Jr. 1988. 'On the Mechanics of Economic Development'. *Journal of Monetary Economics* 22 (1) (July): 3–42.

Michaels, Guy, Ferdinand Rauch, and Stephen J. Redding. 2012. 'Urbanization and Structural Transformation*.' *Quarterly Journal of Economics* 127 (2) (May): 535–586.

Misiewicz, Jolanta K. 1989. 'Positive Definite Norm Dependent Functions on L∞'. *Statistics & Probability Letters* 8 (3): 255–260.

Nagy, David. 2016. *City Location and Economic Development*. 2016 Meeting Papers. Society for Economic Dynamics.

Pedersen, Jan. 2002. 'Periodic Ornstein-Uhlenbeck Processes Driven by Lévy Processes'. *Journal of Applied Probability* 39 (4) (December): 748–763.

Quah, Danny. 2002. 'Spatial Agglomeration Dynamics'. *American Economic Review* 92 (2): 247–252.

Rahman, Qazi I., and Gerhard Schmeisser. 1990. 'Characterization of the Speed of Convergence of the Trapezoidal Rule'. *Numerische Mathematik* 57 (1) (December 1): 123–138.

Redding, Stephen J. 2013. 'Economic Geography: A Review of the Theoretical and Empirical Literature'. In *Palgrave Handbook of International Trade*, edited by Daniel Bernhofen, Rod Falvey, David Greenaway, and Udo Kreickemeier, 497–531. London: Palgrave Macmillan UK.

Rossi-Hansberg, Esteban, and Mark L. J. Wright. 2007. 'Urban Structure and Growth'. *Review of Economic Studies* 74 (2): 597–624.

Schoenberg, Isaac J. 1938. 'Metric Spaces and Positive Definite Functions'. *Transactions of the American Mathematical Society* 44 (3): 522–536.

Appendix A: Continuous stochastic processes

Appendix A.1: Type spaces

We would like to be able to draw realisations of a shock which are continuous (and hence locally correlated) over some compact space X, equipped with a Radon measure. Compactness of X ensures that the law of large numbers does not remove the aggregate impact of the stochastic process, and the existence of a Radon measure is a technical assumption that will not rule out any spaces of interest.

Often, we will have $X = h^{-1}(G)$, where $h: X \to G$ is continuous, and G is a locally-compact, abelian group, also equipped with some Radon measure. For those readers not familiar with group theory, one may view a group as a structure on which "addition" and "subtraction" are defined, along with an identity "zero". The addition and subtraction operations can be thought of as spatial translations. This "additive" group structure will be important, since it will give rise to the spatial analogue of the time series procedure of taking lags or leads.

We will define an underlying continuous stochastic process on G, a realisation of which will be a continuous function $\phi: G \to \mathbb{R}$. It will turn out that continuous stochastic processes will be particularly easy to construct on such groups, explaining our interest in type spaces (Xs), that admit such a representation. The realisation of the stochastic process on X will then be given by $x \mapsto \phi\big(h(x)\big)$, which is continuous by the continuity of ϕ and h. If not stated otherwise, h will be the identity map (or more strictly, the inclusion map). The assumption that G is abelian (i.e. the group operation is commutative) is not needed, but all practical examples will feature abelian G, so nothing is lost. Throughout, the group operation will be denoted "+", with inverse "−" and identity element "0".

For example, we might have:

- $X = [0,1]^n \subseteq \mathbb{R}^n = G$, for some $n \in \mathbb{N}$. This might represent a space of types in which there is a meaningful boundary, such as education levels, or with $n = 2$, the physical area of a country. The group operation and measure are the normal ones on $\mathbb{R}^n$.
- $X = \mathbb{S}_n$, for some $n \in \mathbb{N}$, where $\mathbb{S}_n = \{x \in \mathbb{R}^{n+1} | \|x\| = 1\}$ (i.e. circle, sphere, etc.). This might represent a space of types with no meaningful boundary, in which there is no intrinsic difference between the axes of the type space, or with $n = 2$, it can represent the surface of earth (a sphere!). The group operation on $\mathbb{S}_1$ is addition of angles, so in that case we may take $G = X_1$. Spheres in three-dimensional space do not have an Abelian group operation, so need to be treated specially. The measure on the space is given by the usual spherical surface element.

- $X = \mathbb{T}_n = G$, for some $n \in \mathbb{N}$, where $\mathbb{T}_n = \mathbb{R}^n/\mathbb{Z}^n \simeq \mathbb{S}_1^n$ (i.e. circle, torus, etc.). This gives another representation for type spaces with no boundary, this time for type spaces with clear axes/anisotropy. In this case, the group operation and measure are the ones induced by the quotient construction.
- X is the embedding of a graph in $\mathbb{R}^n$ (i.e. a collection of joined curved line-segments). This might represent a road, river, train or canal network. There is no natural group operation in general, so the existence of continuous processes on this space will be non-trivial. The measure in this case is the one dimensional Hausdorff measure in X.

Appendix A.2: General results on the existence of continuous stochastic processes

Before introducing our general results on the existence of continuous stochastic processes in spaces such as these, we first need to define two closely related notions of positive definiteness for functions.

Definition: Let $f: G \to \mathbb{R}$. We say f is **positive definite on G** if f is a positive, even function, and $\forall n \in \mathbb{N}$, $\forall x_1, \dots, x_n \in G$, the matrix $\left[f\left(x_i - x_j\right)\right]_{i,j=1,\dots,n}$ is positive semi-definite.

Definition: Let $f: X \times X \to \mathbb{R}$. We say f is **positive definite on $X \times X$** if f is a positive, symmetric function, and $\forall n \in \mathbb{N}$, $\forall x_1, \dots, x_n \in G$, the matrix $\left[f\left(x_i, x_j\right)\right]_{i,j=1,\dots,n}$ is positive semi-definite.

The two notions are related, as if f is a positive definite function on G, then $(x_1, x_2) \mapsto f\big(h(x_1) - h(x_2)\big)$ is positive definite on $X \times X$, providing h is as defined previously.

It is easy to verify that sums and products of positive definite functions are positive definite, and that a positive multiple of a positive definite function is positive definite. A further useful characterisation of positive definite functions is Bochner's theorem, which, in our context, implies that f is positive definite on G if and only if the Fourier transform (equivalently, cosine transform) of f on G is positive. For spaces with a group structure, this gives an easy method of constructing positive definite functions.

We are interested in positive definite functions, as by results due to Doob, for the existence of a Gaussian stochastic process with covariance $f\left(x_i, x_j\right)$ for all $x_1, x_2 \in X$, it is sufficient that the function f be positive definite on $X \times X$. Hence, by using this result along with Bochner's theorem, we are easily able to verify if there is a continuous stochastic process with the covariance structure we desire. Alternatively, if we find results in the mathematical literature proving the positive definiteness of some function of interest, then we know there is a continuous stochastic process with that auto-covariance function.

Appendix A.3: Continuous stochastic processes in Euclidean spaces

We now proceed to give examples of spaces and some of the possible continuous stochastic processes on those spaces. Ornstein-Uhlenbeck processes extend naturally to Euclidean spaces, and may be further generalized to allow for different rates of decay of the auto-covariance. In particular, let $\| \quad \|_p$ be the usual ℓ_p (quasi-)norm on $\mathbb{R}^n$, then $x \mapsto \exp\left\{-\|x\|_p^q\right\}$ is positive definite on $\mathbb{R}^n$ if and only if one of the following conditions hold:

- $0 < q \le p \le 2$ (independent of n).
- $n = 1$, and $p > 0$, and $0 < q \le 2$.
- $n = 2$, and $p \in (2, \infty]$, and $q \in (0,1]$,

(Schoenberg 1938; Misiewicz 1989; Koldobsky 1992; see also Kuniewski and Misiewicz 2014). This gives us a wide range of stochastic processes on $\mathbb{R}^n$, with the $p = 1$, $q = 1$ and $p = 2$, $q = 1$ processes both being contenders to be the "natural" Ornstein-Uhlenbeck process on $\mathbb{R}^n$. The $p = q = 2$ process is also potentially useful in macroeconomic applications, as it is the unique stochastic process in this class with realisations that are smooth almost surely.

Another useful class of positive definite functions on $\mathbb{R}^n$ are given by $x \mapsto (1+\|x\|_2^{\alpha})^{-\frac{\tau}{\alpha}}$, where $\alpha \in (0,2]$ and $\tau > 0$ (Gneiting and Schlather 2004). These have long-memory, so may be useful in applications with very high spatial dependence. Further processes on $\mathbb{R}^n$ may be constructed by taking $G = \mathbb{R}^m$, $X = \mathbb{R}^n$, and $h(x) = Ax$ for some matrix A, then $f\big(h(x_1) - h(x_2)\big) = f\big(h(x_1 - x_2)\big)$ so $x \mapsto \exp\left\{-\|Ax\|_p^q\right\}$ and $x \mapsto (1 + \|Ax\|_2^{\alpha})^{-\frac{\tau}{\alpha}}$ are positive definite with the same assumptions as before.

Appendix A.4: Continuous stochastic processes on circles and spheres

If $X = \mathbb{S}_n$, the sphere in $n+1$ dimensional space, then the natural distance between points is the great circle distance, which, appropriately normalised, is given by $d(x_1,x_2) = \frac{1}{2\pi}\arccos(x_1'x_2) \in \left[0,\frac{1}{2}\right]$. Then, for $\zeta > 0$, $(x_1,x_2) \mapsto \exp\{-\zeta d(x_1,x_2)^q\}$ is positive definite on $X \times X$ if and only if $q \in (0,1]$ (Bogomolny, Bohigas, and Schmit 2007). On $\mathbb{S}_1$, this function has cosine transform $k \mapsto \frac{\zeta(1-(-1)^k e^{-\zeta/2})}{4\pi^2k^2+\zeta^2}$ for $k \in \mathbb{N}^+$, which has an undesirable oscillating component not present in the cosine transform on $\mathbb{R}$ of $x \mapsto \exp(-\zeta|x|)$, i.e. $\omega \mapsto \frac{\zeta}{4\pi^2\omega^2+\zeta^2}$. As a result, this may not be a particularly natural choice. As an alternative, it is worth noting that on $\mathbb{S}_1$,[14] for $\zeta > 0$, the function $(x_1,x_2) \mapsto s\big(\zeta, d(x_1,x_2)\big)$ is positive definite, where $s(\zeta,d) = \frac{\exp\left(-\zeta d + \frac{\zeta}{2}\right) + \exp\left(\zeta d - \frac{\zeta}{2}\right)}{\exp\left(\frac{\zeta}{2}\right) + \exp\left(-\frac{\zeta}{2}\right)}$ (Pedersen 2002). Furthermore, this has a cosine transform proportional to $k \mapsto \frac{\zeta}{4\pi^2k^2+\zeta^2}$ for $k \in \mathbb{N}^+$, which means it is the natural generalisation of the Ornstein-Uhlenbeck process on $\mathbb{R}$. Other possibilities include $(x_1,x_2) \mapsto (1-\zeta) + \zeta\left(\frac{1}{2} - d(x_1,x_2)\right)^2$ and $(x_1,x_2) \mapsto (1-\zeta) + \frac{1}{2}\zeta(1 + x_1'x_2)$, which are both positive definite for $\zeta \in [0,1]$, by the condition given in Gneiting (2013). The latter is an analogue of $x \mapsto \exp(-\zeta x^2)$ on $\mathbb{R}$, and will lead to smooth sample paths.

Appendix A.5: Continuous stochastic processes on tori

If $G = G_1 \times G_2 \times \ldots \times G_n$, and f_i is positive definite on G_i for $i = 1, \ldots, n$, then the function $f\colon G \to \mathbb{R}$ defined by $f(x) = \prod_{i=1}^n f_i(x_i)$ is positive definite on G. Hence, positive definite functions on tori can be constructed from products of positive definite functions on circles (where the circle is identified with $\mathbb{R}/\mathbb{Z}$). For example, by our previous results, for $\zeta_1, \ldots, \zeta_n > 0$, $q_1, \ldots, q_n \in (0,1]$, the function:

$$(x_1, \ldots, x_n) \mapsto \exp\left\{-\sum_{i=1}^n \zeta_i \min\{|x_i|, 1-|x_i|\}^{q_i}\right\}$$

will be positive definite on G.

Appendix A.6: Continuous stochastic processes on graphs or networks

In general, graphs cannot be isometrically embedded in Euclidean space, so if d is the shortest path metric on the (embedding of the) graph, $(x_1,x_2) \mapsto \exp\big(-\zeta d(x_1,x_2)^2\big)$ will only be positive definite for very particular graphs. We do know however that $(x_1,x_2) \mapsto \exp\big(-\zeta d(x_1,x_2)\big)$ will be positive definite for all $\zeta > 0$ if the following conditions are all satisfied (Chepoi, Deza, and Grishukhin 1997):

- the graph is unweighted, or possesses integer weights,
- the graph is planar,
- every interior face of the graph is an isometric cycle,

[14] This does not hold on $\mathbb{S}_2$ or higher, by the condition given in Gneiting (2013).

- two interior faces meet at at most one edge (of length one).

Additionally, $(x_1, x_2) \mapsto \exp(-\zeta d(x_1, x_2))$ is positive definite on all (weighted) trees, for all $\zeta > 0$ (Hjorth et al. 1998).

Furthermore, it may be shown that for all graphs, there exists $\bar{\zeta}$ such that $(x_1, x_2) \mapsto \exp(-\zeta d(x_1, x_2))$ is positive definite for $\zeta > \bar{\zeta}$. The idea of the proof is the following. For large ζ, the "ACF" matrix is diagonally dominant on the graph's vertices, hence positive definite, hence we can draw from a finite dimensional Gaussian process on the vertices, and then link the realisations at the vertices with independent Ornstein-Uhlenbeck processes along each, conditional on them taking the given values at the vertices.

Appendix B: Further model properties

Appendix B.1: Lagrangians

The capital holding company's problem leads to the following Lagrangian:

$$\mathbb{E}_t \sum_{s=0}^{\infty} \left[\prod_{k=1}^{s} \Xi_{t+k} \right] \left(\mathcal{R}_{K,x,t+s} K_{x,t+s-1} - P_{x,t} I_{x,t+s} + Q_{x,t} \left((1-\delta_K) K_{x,t-1} + \left[1 - \Phi\left(\frac{I_{x,t}}{I_{x,t-1}} \right) \right] I_{x,t} - K_{x,t} \right) \right).$$

The household's problem leads to the following Lagrangian:

$$\begin{aligned}
&\mathbb{E}_t \sum_{s=0}^{\infty} \left[\prod_{k=1}^{s} \beta_{t+k-1} \right] \left[\int_X \left[N_{x,t+s-1} \left[\frac{U_{x,t+s}^{1-\varsigma}}{1-\varsigma} \right.\right.\right. \\
&\qquad + \mu_{U,x,t+s} \left[\left(\frac{C_{x,t+s}}{N_{x,t+s-1}} \right)^{\theta_C} \left(\frac{E_{x,t+s}}{N_{x,t+s-1}} \right)^{\theta_F} \left(\frac{1-L_{x,t+s}}{N_{x,t+s-1}} \right)^{\theta_L} \left(\frac{1}{1+\nu} \Gamma^{1+\nu} - \frac{1}{1+\nu} \left(\frac{H_{x,t+s}}{N_{x,t+s-1}} \right)^{1+\nu} \right)^{\theta_H} \right. \\
&\qquad \cdot \left(\frac{1}{2} \Omega^2 - \frac{1}{2} \left(\log\left(\frac{N_{x,t+s-1}}{N_{t+s-1}} \right) \right)^2 \right)^{\theta_N} \\
&\qquad \cdot \left(1 - \frac{\int_X \mathcal{N}_{x,\tilde{x},t+s} \, d\tilde{x}}{N_{x,t+s-1}} \right)^{\psi_1} \left(\bar{d} - \frac{\int_X d(x,\tilde{x}) \mathcal{N}_{x,\tilde{x},t+s} \, d\tilde{x}}{\int_X \mathcal{N}_{x,\tilde{x},t+s} \, d\tilde{x}} \right)^{\psi_2} \exp\left[\psi_3 \frac{\int_X N_{\tilde{x},t+s-1} \log \frac{\mathcal{N}_{x,\tilde{x},t+s}}{N_{x,t+s-1}} \, d\tilde{x}}{\int_X N_{\tilde{x},t+s-1} \, d\tilde{x}} \right] \\
&\qquad \left.\left. - U_{x,t+s} \right]\right] + \frac{\mu_{N,x,t+s}}{1-\varsigma} \left[G_{N,t+s} N_{x,t+s-1} - \int_X \mathcal{N}_{x,\tilde{x},t+s} \, d\tilde{x} + \int_X \mathcal{N}_{\tilde{x},x,t+s} \, d\tilde{x} - N_{x,t+s} \right] \Bigg] \, dx \\
&\qquad + \mu_{B,t+s} \left[\int_X (\mathcal{R}_{L,x,t+s} L_{x,t+s} + W_{x,t+s} H_{x,t+s}) \, dx + R_{t+s-1} B_{t+s-1} + \mathrm{T}_{t+s} \right. \\
&\qquad \left.\left. - \int_X (P_{x,t+s} C_{x,t+s} + E_{x,t+s}) \, dx - B_{t+s} \right] \right].
\end{aligned}$$

Note that $\mu_{U,x,t}$ and $\mu_{B,t}$ do not occur in the first order conditions given in the text, as we substitute it out from the first order condition for $U_{x,t}$:

$$\mu_{U,x,t} = U_{x,t}^{-\varsigma}.$$

and the first order condition of $E_{x,t}$:

$$\mu_{B,t} = \theta_F \frac{N_{x,t-1}}{E_{x,t}} U_{x,t}^{1-\varsigma}.$$

Appendix B.2: Equilibrium conditions

The complete set of equilibrium conditions of the model are as follows:

$$\overline{Y}_{x,t} := \int_X Y_{\tilde{x},t} P_{\tilde{x},t}^{\frac{1+\lambda}{\lambda}} \exp\left[-\frac{\tau_t}{\lambda} d(\tilde{x}, x)\right] d\tilde{x}$$

$$\Pi_{x,t} := \frac{\lambda}{1+\lambda} (1+\lambda)^{-\frac{1}{\lambda}} \mathcal{P}_{x,t}^{-\frac{1}{\lambda}} \overline{Y}_{x,t}$$

$$P_{x,t} = (1+\lambda) \left[\int_X J_{\tilde{x},t} (\mathcal{P}_{x,t} \exp[\tau_t d(x,\tilde{x})])^{-\frac{1}{\lambda}} d\tilde{x}\right]^{-\lambda}$$

$$\phi_t \mathcal{P}_{x,t} = \Pi_{x,t} + (1-\delta_J) \mathbb{E}_t \Xi_{t+1} \phi_{t+1} \mathcal{P}_{x,t+1}$$

$$F_{x,t} = L_{x,t}^{\gamma} Z_{F,x,t}^{1-\gamma}$$

$$\gamma \frac{F_{x,t}}{L_{x,t}} = \mathcal{R}_{L,x,t}$$

$$(1-\gamma) \frac{F_{x,t}}{Z_{F,x,t}} = \mathcal{P}_{x,t}$$

$$Z_{x,t} = \left[K_{x,t-1}^{\alpha} (A_{x,t} H_{x,t})^{1-\alpha}\right]^{1-\kappa} M_{x,t}^{\kappa}$$

$$(1-\kappa)\alpha \mathcal{P}_{x,t} \frac{Z_{x,t}}{K_{x,t-1}} = \mathcal{R}_{K,x,t}$$

$$(1-\kappa)(1-\alpha) \mathcal{P}_{x,t} \frac{Z_{x,t}}{H_{x,t}} = W_{x,t}$$

$$\kappa \mathcal{P}_{x,t} \frac{Z_{x,t}}{M_{x,t}} = P_{x,t}$$

$$K_{x,t} = (1-\delta_K) K_{x,t-1} + \left[1 - \Phi\left(\frac{I_{x,t}}{I_{x,t-1}}\right)\right] I_{x,t}$$

$$1 = \mathbb{E}_t \Xi_{t+1} \frac{\mathcal{R}_{K,x,t+1} + Q_{x,t+1}(1-\delta_K)}{Q_{x,t}}$$

$$P_{x,t} = Q_{x,t}\left(1 - \Phi\left(\frac{I_{x,t}}{I_{x,t-1}}\right) - \Phi'\left(\frac{I_{x,t}}{I_{x,t-1}}\right)\frac{I_{x,t}}{I_{x,t-1}}\right) + \mathbb{E}_t \Xi_{t+1} Q_{x,t+1} \Phi'\left(\frac{I_{x,t+1}}{I_{x,t}}\right)\left(\frac{I_{x,t+1}}{I_{x,t}}\right)^2$$

$$Y_{x,t} = C_{x,t} + I_{x,t} + M_{x,t}$$

$$Z_{x,t} = Z_{F,x,t} + \phi_t [J_{x,t} - (1-\delta_J) J_{x,t-1}] + J_{x,t} (1+\lambda)^{-\frac{1+\lambda}{\lambda}} \mathcal{P}_{x,t}^{-\frac{1+\lambda}{\lambda}} \overline{Y}_{x,t}$$

$$\int_X E_{x,t}\, dx = \int_X F_{x,t}\, dx$$

$$U_{x,t} = \left(\frac{C_{x,t}}{N_{x,t-1}}\right)^{\theta_C} \left(\frac{E_{x,t}}{N_{x,t-1}}\right)^{\theta_F} \left(\frac{1-L_{x,t}}{N_{x,t-1}}\right)^{\theta_L} \left(\frac{1}{1+\nu}\Gamma^{1+\nu} - \frac{1}{1+\nu}\left(\frac{H_{x,t}}{N_{x,t-1}}\right)^{1+\nu}\right)^{\theta_H} \left(\Omega - \frac{1}{2}\left(\log\left(\frac{N_{x,t-1}}{N_{t-1}}\right)\right)^2\right)^{\theta_N}$$

$$\cdot \left(1 - \frac{\mathcal{N}_{x,t}}{N_{x,t-1}}\right)^{\psi_1} \left(\bar{d} - \frac{\mathcal{D}_{x,t}}{\mathcal{N}_{x,t}}\right)^{\psi_2} \exp\left[\psi_3 \int_X \frac{N_{\tilde{x},t-1}}{N_{t-1}} \log \frac{\mathcal{N}_{x,\tilde{x},t}}{N_{x,t-1}} d\tilde{x}\right],$$

$$N_t := \int_X N_{\tilde{x},t}\, d\tilde{x},$$

$$\mathcal{N}_{x,t} := \int_X \mathcal{N}_{x,\tilde{x},t}\, d\tilde{x},$$

$$\mathcal{D}_{x,t} := \int_X d(x,\tilde{x}) \mathcal{N}_{x,\tilde{x},t}\, d\tilde{x}$$

$$N_{x,t} = G_{N,t} N_{x,t-1} - \mathcal{N}_{x,t} + \int_X \mathcal{N}_{\tilde{x},x,t}\, d\tilde{x}$$

$$\theta_C E_{x,t} = \theta_F P_{x,t} C_{x,t}$$

$$\theta_L E_{x,t} = \theta_F \mathcal{R}_{L,x,t} (1 - L_{x,t})$$

$$\theta_H \left(\frac{H_{x,t}}{N_{x,t-1}}\right)^{\nu} = \theta_F \frac{N_{x,t-1}}{E_{x,t}} W_{x,t} \left(\frac{1}{1+\nu}\Gamma^{1+\nu} - \frac{1}{1+\nu}\left(\frac{H_{x,t}}{N_{x,t-1}}\right)^{1+\nu}\right)$$

$$1 = R_t \mathbb{E} \Xi_{t+1}$$

$$\Xi_{t+1} = \beta_t \frac{N_{x,t} E_{x,t} U_{x,t+1}^{1-\varsigma}}{E_{x,t+1} N_{x,t-1} U_{x,t}^{1-\varsigma}}$$

$$\begin{aligned}\mu_{N,x,t} = \beta_t \mathbb{E}_t \Bigg[& \mu_{N,x,t+1} G_{N,t+1} + U_{x,t+1}^{1-\varsigma} \\ & + (1-\varsigma) U_{x,t+1}^{1-\varsigma} \Bigg[\theta_H \frac{\left(\frac{H_{x,t+1}}{N_{x,t}}\right)^{1+\nu}}{\frac{1}{1+\nu}\Gamma^{1+\nu} - \frac{1}{1+\nu}\left(\frac{H_{x,t+1}}{N_{x,t}}\right)^{1+\nu}} - \theta_N \frac{\log\left(\frac{N_{x,t}}{N_t}\right)}{\frac{1}{2}\Omega^2 - \frac{1}{2}\left(\log\left(\frac{N_{x,t}}{N_t}\right)\right)^2} + \psi_1 \frac{\mathcal{N}_{x,t+1}}{N_{x,t} - \mathcal{N}_{x,t+1}} \\ & - (\theta_C + \theta_F + \theta_L + \psi_3) \Bigg] \Bigg]\end{aligned}$$

$$\mu_{N,x,t} = \mu_{N,\check{x},t} + (1-\varsigma) N_{x,t-1} U_{x,t}^{1-\varsigma} \left[\psi_3 \frac{N_{\check{x},t-1}}{N_{t-1}\mathcal{N}_{x,\check{x},t}} - \psi_1 \frac{1}{N_{x,t-1} - \mathcal{N}_{x,t}} - \psi_2 \frac{d(x,\check{x})\mathcal{N}_{x,t} - \mathcal{D}_{x,t}}{\bar{d}\mathcal{N}_{x,t}^2 - \mathcal{N}_{x,t}\mathcal{D}_{x,t}} \right]$$