Supplementary Material
Hinterlands, city formation and growth: Evidence from the U.S. westward expansion

Dávid Krisztián Nagy*

February 23, 2023

A Theory appendix

This appendix supplements the theoretical framework presented in Section 3 in six respects. First, Section A.1 defines the equilibrium of the multi-country model. Next, Section A.2 shows how the population distribution in period t in the multi-country model can be obtained by solving a system of three equations, and describes an algorithm to solve these equations. Section A.3 presents a special case of the model with a homogenous hinterland around a non-farm location. Section A.4 discusses why multiplicity of equilibria may arise in the model, as well as why the complex structure of the model does not allow for a theoretical characterization of equilibrium uniqueness. Section A.5 presents a version of the model in which consumption happens at the residential location, as well as a condition on shipping costs that guarantees isomorphism between this alternative model and the model of Section 3. Finally, Section A.6 provides proofs.

A.1 Definition of the equilibrium in the multi-country model

Definition. *Given parameters α, β, γ, δ, ϵ, λ, μ, ν, f, the set of locations S, countries' population $\bar{L}_{ct}$ and growth shifters f_{ct}, as well as functions $H_t, A_1, B : S \to \mathbb{R}$ and $\tau_t^F, \tau_t^N : S^2 \to \mathbb{R}$, an* ***equilibrium*** *of the economy is a set of functions $L_t^F, L_t^N, p_t^F, p_t^N, q_t^N, h_t, \ell_t^P, \ell_t^I, x_t^F, \bar{c}_t, n_t, P_t, w_t, R_t, A_t : S \to \mathbb{R}$ and $\sigma_t : S \to S$, as well as countries' utility levels U_{ct} for each time period $t \in \{1, 2, ...\}$ such that the following hold:*

1. *Farmers maximize utility, implying that farmers living at r choose their trading place $\sigma_t(r)$ such that*

$$\sigma_t(r) = argmax_{s \in S_c} U_t^F(r, s)$$

*CREi, Universitat Pompeu Fabra and BSE. Email: dnagy@crei.cat.

and their utility is equalized across locations within each country, that is,

$$U_t^F(r, \sigma_t(r)) = U_t^F(u, \sigma_t(u)) \tag{A.1}$$

for any $r, u \in S_c$*, such that* $U_t^F(\cdot, \cdot)$ *is given by (3.1).*

2. *Non-farm workers maximize utility, and their utility is equalized across locations within each country, that is,*

$$U_t^N(s) = U_t^N(u) \tag{A.2}$$

for any $s, u \in S_c$ *with positive non-farm population, where* $U_t^N(\cdot)$ *is given by (3.2).*

3. *Consumers' utility is equalized across sectors within each country. Hence,*

$$U_{ct} = U_t^F(r, s) = U_t^N(s) \tag{A.3}$$

for any $r, s \in S_c$ *such that* $s = \sigma_t(r)$.

4. *Non-farm firms maximize profits, and free entry drives down profits to zero. Hence, their price is given by (3.5), their output by (3.6), their unit cost per wage by (3.7), and their factor use by (3.8) to (3.10). The number of goods produced at location* s *is given by (3.11), and the price index at* s *is*

$$P_t(s) = \left[\int_S n_t(u) p_t^N(u)^{1-\epsilon} \tau_t^N(u, s)^{1-\epsilon} du\right]^{\frac{1}{1-\epsilon}}. \tag{A.4}$$

5. *The market for the farm good clears at each trading location* s*. That is,*

$$\int_{\sigma_t^{-1}(s)} \nu \tau_t^F(r, s)^{-1} B(r) L_t^F(r)^{\alpha} H_t(r)^{1-\alpha} dr = x_t^F(s) n_t(s) + (1-\nu) \frac{w_t(s)}{p_t^F(s)} L_t^N(s) \tag{A.5}$$

where $\sigma_t^{-1}(s)$ *denotes the set of locations from which farmers ship to* s*, the left-hand side corresponds to supply net of farmers' demand at* s*, the first term on the right-hand side corresponds to firms' demand, and the second term on the right-hand side corresponds to non-farm workers' demand.*

6. *The market for non-farm goods clears at each trading location* s*. That is,*

$$q_t^N(s) = \int_S \nu p_t^N(s)^{-\epsilon} \tau_t^N(s, u)^{1-\epsilon} P_t(u)^{\epsilon-1} I_t(u) du \tag{A.6}$$

where the left-hand side corresponds to the supply of any non-farm variety produced at s*, and the right-hand side corresponds to total demand for the variety.* $I_t(u)$ *denotes total income of consumers at* u*, and equals the sum of farmers' and non-farm workers'* *income:*

$$I_t(u) = \int_{\sigma_t^{-1}(u)} p_t^F(u)\,\tau_t^F(r,u)^{-1}\,B(r)\,L_t^F(r)^{\alpha}\,H_t(r)^{1-\alpha}\,dr + w_t(u)\,L_t^N(u)$$

7. *National labor markets clear. That is,*

$$\bar{L}_{ct} = \int_{S_c} \left[L_t^F(r) + L_t^N(r)\right] dr \tag{A.7}$$

where S_c denotes the set of locations that belong to country c.

8. *The market for land clears at each location. That is,*

$$H_t(r) = L_t^F(r)\,h_t(r)\,. \tag{A.8}$$

9. *Productivity levels evolve according to equations (3.20) and (3.4).*

The equilibrium of the multi-country model differs in three respects from the equilibrium of the single-country model, which was defined in Section 3.1.7. First, as consumers cannot move across countries, utility equalization conditions (A.1), (A.2) and (A.3) are only imposed across locations that are in the same country. Second, the national labor market clearing condition (A.7) is imposed by country, rather than across all locations. Finally, equation (3.20) replaces equation (3.3) as the law of motion for local productivity.

A.2 Solving the equilibrium population distribution in period t

To simplify the model's period-t equilibrium conditions – that is, the system of equations (3.1), (3.2), (3.5) to (3.11) and (A.1) to (A.8) –, note first that farmers living at a trading place always want to trade there. The intuition for this result is that, if farmers living at s preferred some other trading place u to s, then, by the triangle inequality of shipping costs, farmers living at any other location would also prefer u to s. Hence, s would not even arise as a trading place.

But what is the trading place chosen by farmers who do not live at one? Clearly, farmers living at a location r in country c choose the trading place that maximizes their utility (3.1). Note also that, by (A.3), the utility of a farmer living at a trading place $s \in S_c$ equals the utility of a non-farmer (3.2),

$$\frac{p_t^F(s)\,B(s)\left[\frac{H_t(s)}{L_t^F(s)}\right]^{1-\alpha}}{P_t(s)^{\nu}\,p_t^F(s)^{1-\nu}} = \frac{w_t(s)}{P_t(s)^{\nu}\,p_t^F(s)^{1-\nu}}$$

from which the price of the farm good can be expressed as

$$p_t^F(s) = B(s)^{-1}\left[\frac{L_t^F(s)}{H_t(s)}\right]^{1-\alpha} w_t(s). \tag{A.9}$$

Plugging this back into (3.1), one obtains

$$U_t^F(r,s) = \tau_t^F(r,s)^{-1}\frac{B(r)}{B(s)}\left[\frac{H_t(r)}{H_t(s)}\right]^{1-\alpha}\left[\frac{L_t^F(r)}{L_t^F(s)}\right]^{-(1-\alpha)} U_{ct}$$

where I used (A.3) again to substitute $U_{ct} = \frac{w_t(s)}{P_t(s)^{\nu} p_t^F(s)^{1-\nu}}$. The trading place s which is optimal for farmers at location r is the one that maximizes the above expression, thus it is

$$\sigma_t(r) = argmax_{s\in S_c}\tau_t^F(r,s)^{-1} B(s)^{-1} H(s)^{-(1-\alpha)} L_t^F(s)^{1-\alpha}. \tag{A.10}$$

Once we know who trades where, utility equalization relates the farm population of any location to the farm population of its trading place. To see this, consider a location r together with its trading place $\sigma_t(r)$. By (A.1), farmers living at these two places have the same utility, that is,

$$U_t^F(r,\sigma_t(r)) = U_t^F(\sigma_t(r),\sigma_t(r)).$$

Substituting for $U_t^F(\cdot,\cdot)$ using (3.1), one can express the farm population of r as

$$L_t^F(r) = \tau_t^F(r,\sigma_t(r))^{-\frac{1}{1-\alpha}}\frac{H_t(r)}{H_t(\sigma_t(r))}\left[\frac{B(r)}{B(\sigma_t(r))}\right]^{\frac{1}{1-\alpha}} L_t^F(\sigma_t(r)) \tag{A.11}$$

We are only left with finding the distribution of farmers across trading places since equations (A.10) and (A.11) pin down farm population at any location r *conditional* on this distribution. To obtain the farm population of trading places, I use the price index (A.4), the non-farm market clearing condition (A.6), and utility equalization. The result is stated in the following lemma.

Lemma 2. *In any period t, the distribution of farm population is the solution to the following system of equations:*

$$B(s)^{-\frac{1-\nu}{\nu}\frac{\epsilon(\epsilon-1)}{2\epsilon-1}}\left[\frac{L_t^F(s)}{H_t(s)}\right]^{(1-\alpha)\frac{1-\nu}{\nu}\frac{\epsilon(\epsilon-1)}{2\epsilon-1}}\bar{c}_t(s)^{\frac{(\epsilon-1)^2}{(2\epsilon-1)}} = \kappa U_{c(s),t}^{-\frac{\epsilon(\epsilon-1)}{2\epsilon-1}}.$$

$$\int_S U_{c(u),t}^{-\frac{(\epsilon-1)^2}{2\epsilon-1}}\frac{B(u)^{\frac{1-\nu}{\nu}\frac{(\epsilon-1)^2}{2\epsilon-1}}\left[\frac{L_t^F(u)}{H_t(u)}\right]^{-(1-\alpha)\frac{1-\nu}{\nu}\frac{(\epsilon-1)^2}{2\epsilon-1}}\bar{c}_t(u)^{-\frac{\epsilon(\epsilon-1)}{(2\epsilon-1)}}}{(1-\nu)\left[\epsilon(1-\mu)+\mu\right]+(\epsilon-1)\mu\left[1-\nu\bar{c}_t(u)^{\beta-1}\right]}\left[\int_{\sigma_t^{-1}(u)} L_t^F(r)\,dr\right]\tau_t^N(u,s)^{1-\epsilon}\,du \tag{A.12}$$

where $L_t^F(r)$ as a function of $L_t^F(u)$ is given by (A.11), $\sigma_t^{-1}(u)$ is given by (A.10), $\bar{c}_t(s)$ is given by (3.7), and $\kappa = \epsilon^{1-\epsilon}(\epsilon-1)^{\epsilon-2}\mu^{-1}f^{-1}$ is a constant.

Proof. See Appendix A.6.

□

Next, non-farm population at trading places can be obtained using the distribution of farm population and farm market clearing (A.5). Combining (A.5) with (3.10) and (3.11) yields

$$L_t^N(s) = \frac{\nu}{(1-\nu)+\mu(\epsilon-1)\left[\mu+\epsilon(1-\mu)+\frac{\epsilon}{\bar{c}_t(s)^{1-\beta}-1}\right]^{-1}} \int_{\sigma_t^{-1}(s)} L_t^F(r)\,dr \qquad \text{(A.13)}$$

as the non-farm population of location s. Note that, by (A.13), the non-farm population of places to which no one ships farm goods is zero. Since the supply of farm goods is zero at these locations, non-farm production cannot take place, hence non-farm workers do not move to these locations in equilibrium. On the other hand, equation (A.11) implies that every location with a positive amount of land has some farm population. Pulling these results together, one can conclude that there are two types of locations in equilibrium: locations that have both farm and non-farm population and serve as trading places for farmers, and locations that only have farm population and do not serve as trading places. Whether a given location belongs to the first or the second type is an endogenous outcome shaped by the fundamentals of the model, including geography.

Finally, countries' utility levels U_{ct} can be obtained by imposing national labor market conditions (A.7). Hence, we have the following lemma.

Lemma 3. *Given the values of parameters, geography S, countries' population $\bar{L}_{ct}$ and growth shifters f_{ct}, functions $H_t(\cdot)$, $B(\cdot)$, $\tau_t^F(\cdot,\cdot)$ and $\tau_t^N(\cdot,\cdot)$ and the distribution of non-farm productivity $A_t(\cdot)$, the system of three equations (A.7), (A.12) and (A.13) determines the spatial distribution of population and countries' utility levels in period t.*

As a result of Lemma 3, obtaining the population distribution in period t requires solving the system of equations (A.7), (A.12) and (A.13). In the two-country world considered in this paper, I solve this system by an iteration algorithm, similar to the one applied in Nagy (2022). The algorithm consists of the following steps.

1. Calculate the non-farm productivity of Europe, using equation (4.1).

2. Guess country utility levels $U_{US,t}$ and U_{et}.

3. Guess the farm population trading at each location, $\int_{\sigma_t^{-1}(s)} L_t^F(r)\,dr$.

4. Guess the farm population living at each trading location, $L_t^F(s)$.

5. Use equation (3.7) to calculate $\bar{c}_t(s)$. Then calculate the right-hand side of equation (A.12). Setting the left-hand side equal to the right-hand side, update $\bar{c}_t(s)$. Then use equation (3.7) again to update $L_t^F(s)$. Keep updating $L_t^F(s)$ until convergence.

6. Use the values of $L_t^F(s)$, as well as equations (A.10) and (A.11) to calculate optimal trading places and the farm population of any location r. Using these, update the farm population trading at each location, $\int_{\sigma_t^{-1}(s)} L_t^F(r)\,dr$, and continue from step 4. Keep updating $\int_{\sigma_t^{-1}(s)} L_t^F(r)\,dr$ until convergence.

7. Solve equation (A.13) for non-farm populations $L_t^N(s)$.

8. Check if national labor market clearing conditions (A.7) hold. If not, modify $U_{US,t}$ and U_{et}, and continue from step 3.

Although the complex structure of the model does not allow me to derive conditions under which the algorithm converges to the equilibrium distribution of population, simulation results suggest that the algorithm displays good convergence properties unless either agglomeration or dispersion forces are very strong. In particular, the algorithm always converges to the equilibrium in a broad neighborhood around the parameter values chosen in the calibration.

A.3 A special case of the model with a homogenous hinterland

This appendix section presents a special case of the model in which I can explicitly characterize how changes in geography around a location (in the form of railroad construction, an expansion of political borders, or a decrease in international trade costs) affect the size of the location's effective hinterland, $\tilde{H}_t(s)$.

The key simplification relative to the general model of Section 3 is that I assume a geographically homogenous hinterland around location s. Specifically, I assume that in a Borel measurable set of locations around s, denoted by $S_t(s) \subset S$ such that $s \in S_t(s)$,[1]

- farm productivity is homogenous: for any location $r \in S_t(s)$, $B(r) = B(s)$;
- non-farm productivity is zero everywhere except at s, guaranteeing that no other location in $S_t(s)$ specializes in non-farm production;
- farm trade costs are uniform: between s and any location $r \in S_t(s)$, $\tau_t^F(r,s) = \tau_t(s)$.

I also assume that farm trade costs between locations in $S_t(s)$ and locations in the rest of the economy are infinitely high. This guarantees that the locations in $S_t(s)$, and only those locations, serve as the farm hinterland of s.

[1]With a slight abuse of notation, I also use $S_t(s)$ to denote the mass of locations in set $S_t(s)$.

As for non-farm trade costs, I assume that s can trade with another location e, which represents foreign markets and looks identical to s in every respect. This means that, among other things, e is also surrounded by a hinterland identical to $S_t(s)$. Trading non-farm goods between s and e is subject to an international trade cost $\tau_t^N(s,e) = b_t(s)$. Trading non-farm goods between s or e and any other location in the economy is subject to infinitely high costs. Thus, s and e ony trade non-farm goods with each other in equilibrium.

I need to impose a few further simplifications to keep analytical tractability. First, I need to assume that the share of farm goods in consumption is zero ($\nu = 1$), implying that farm goods are only used as intermediates in the non-farm sector. While a farm good share different from zero might be important for the quantitative results of this paper, the key mechanism of the model hinges on the use of farm goods in non-farm production (see the discussion in Section 3.3), which is present even when $\nu = 1$. Second, I need to assume that the economy-wide utility level, $\bar{U}_t$, is unaffected by what happens to the locations in $S_t(s)$. This assumption can be interpreted as $S_t(s)$ being small relative to the entire economy. Finally, I need to restrict the non-farm productivity of s to be low enough initially. In particular, I need to assume that $A_t(s)$ is such that non-farm firms' unit costs satisfy

$$\bar{c}_t(s)^{1-\beta} - 1 > \frac{\frac{1}{(1-\alpha)\mu} - (1-\beta)}{\epsilon - 1 - \frac{1}{(1-\alpha)\mu}}.$$

This means that s cannot be too developed in the beginning. This assumption suits locations in the U.S. West well in my historical context.

Under these assumptions, the next proposition shows that the size of s's effective hinterland increases when railroad construction, an expansion of political borders, or a fall in international trade costs affects s and its hinterland.

Proposition 2. *In the model with a homogenous hinterland, the size of the effective hinterland of location s, $\tilde{H}_t(s)$, increases if*

1. *railroad construction lowers the cost of shipping farm goods between s and its hinterland, $\tau_t(s)$, or*
2. *political border changes increase the number of locations in s's hinterland, $S_t(s)$, or*
3. *the cost of international trade between s and the rest of the world, $b_t(s)$, decreases.*

Proof. See Appendix A.6.

□

Once Proposition 2 holds, Proposition 1 implies that changes in geography in the form of railroad construction, border expansion, or increasing international trade bring more farm workers to the hinterland of s, and more non-farm workers and innovation to s itself.

In other words, these changes in geography foster farm settlement, city development and growth around s. This result is stated in the following corollary.

Corollary of Proposition 2. *In the model with a homogenous hinterland, any of the three changes in geographic fundamentals presented in parts 1 to 3 of Proposition 2 raises*

i. the farm population of any location in s's hinterland $r \in \sigma_t^{-1}(s)$*, and thus the total farm population of the hinterland,*

ii. the non-farm population of s, and thus the likelihood that s is a city,

iii. innovation at s, and thus total innovation and growth in the economy.

Proof. This is the immediate consequence of Proposition 2 and the Corollary of Proposition 1.

□

A.4 Difficulties of characterizing equilibrium uniqueness

This section of the appendix builds intuition for why it is difficult to theoretically characterize the uniqueness of the period-t equilibrium in the model of Section 3. To this end, I consider two special cases of the model in which uniqueness can be characterized. Next, I turn to the general case and point out how features of the model that are not present in the special cases prevent the theoretical characterization of uniqueness. For simplicity, I restrict my attention to the single-country version of the model in this section of the appendix. Generalizing the results to the multi-country version is straightforward.

Case 1: An economy with only a farm sector. I first consider a special case of the model in which the share of non-farm goods in consumption is zero: $\nu = 0$. This implies that there is no demand for non-farm goods in the economy. Therefore, the non-farm sector is absent and does not employ any workers, implying that the population of any location r equals the location's farm population: $L_t(r) = L_t^F(r)$.

As the farm sector produces a homogeneous good under constant returns to scale, no equilibrium of this economy can feature trade by the spatial impossibility theorem (Starrett, 1978). As locations are in autarky, the utility that a farmer residing at r obtains is given by

$$U_t(r) = x_t^F(r) = B(r)\left[\frac{H_t(r)}{L_t(r)}\right]^{1-\alpha}.$$

By free mobility, $U_t(r)$ becomes equal to the utility level of the country at each location: $U_t(r) = U_t$. Substituting this into the previous equation, one can express the equilibrium population of location r at time t as

$$L_t(r) = U_t^{-\frac{1}{1-\alpha}} B(r)^{\frac{1}{1-\alpha}} H_t(r) \tag{A.14}$$

while the condition $\bar{L}_t = \int_S L_t(r)\, dr$ pins down the level of utility as

$$U_t = \bar{L}_t^{-(1-\alpha)} \left[\int_S B(r)^{\frac{1}{1-\alpha}} H_t(r)\, dr \right]^{1-\alpha}. \tag{A.15}$$

Equations (A.14) and (A.15) uniquely pin down the population of every location. Thus, in the version of the model with only a farm sector, the equilibrium is guaranteed to be unique. This is not surprising as the farm sector does not feature any agglomeration force (such as economies of scale), while it features a force of dispersion in the form of land use. If more people move to a location, the share of land that any of them can use decreases, decreasing farm output and hence utility per individual. This dispersion force prevents the emergence of multiple equilibria.

Case 2: An economy with only a non-farm sector. I now consider a special case of the model in which the share of farm goods in consumption is zero, $\nu = 1$, and farm goods are not used as intermediates in non-farm production either, $\mu = 0$. As a result, there is no demand for farm goods and farmers, and the population of any location r equals the location's non-farm population: $L_t(r) = L_t^N(r)$.

Plugging $\nu = 1$ and $\mu = 0$ into the period-t equilibrium conditions of the model, I obtain that the spatial distribution of population at time t is the solution to the system of equations

$$\bar{c}_t(s)^{\frac{(\epsilon-1)^2}{2\epsilon-1}} = \frac{\tilde{\kappa}}{\epsilon} U_t^{1-\epsilon} \int_S \bar{c}_t(u)^{-\frac{\epsilon(\epsilon-1)}{2\epsilon-1}} L_t(u)\, \tau_t^N(s,u)^{1-\epsilon}\, du \tag{A.16}$$

where $\tilde{\kappa} = \left(\frac{\epsilon-1}{\epsilon}\right)^{\epsilon-1} f^{-1}$, $\bar{c}_t(s) = \left[1 + A_t(s)^{\beta-1}\right]^{\frac{1}{1-\beta}}$, and the level of utility, U_t, is pinned down by the national labor market clearing condition

$$\bar{L} = \int_S L_t(r)\, dr.$$

Equation (A.16) is a special case of the population equations considered in Allen and Arkolakis (2014), which take the form

$$f_1(s)\, L_t(s)^{\gamma_1} = k U_t^{1-\epsilon} \int_S f_2(u)\, L_t(u)^{\gamma_2}\, du \tag{A.17}$$

such that $f_1(s)$ and $f_2(u)$ are exogenously given continuous functions, and k is a constant. Theorem 2 in Allen and Arkolakis (2014) shows that equation (A.17) has a unique positive solution for the population distribution, $L_t(\cdot)$, if $\frac{\gamma_2}{\gamma_1} \in [-1, 1]$. This condition does not hold for equation (A.16), as $\gamma_1 = 0$ and $\gamma_2 = 1$ in that case. Although the condition $\frac{\gamma_2}{\gamma_1} \in [-1, 1]$ is not necessary for uniqueness, Allen and Arkolakis (2014) show that one can construct geographies featuring multiple equilibria if it does not hold. Hence, one can conclude that the version of the model with only a non-farm sector tends to feature multiple equilibria in general.[2]

The fact that an economy with only a non-farm sector tends to feature equilibrium multiplicity is not surprising either. The non-farm sector itself does not feature any dispersion force such as land use, but it features the classical agglomeration force of Krugman (1991), due to transport costs and economies of scale. If more people move to a location, more firms enter the location's non-farm sector, which decreases the price index of non-farm goods. Moreover, local non-farm firms face a larger demand for their product, which allows them to pay higher nominal wages. Higher nominal wages and lower non-farm prices, in turn, attract even more people to the location. This force of circular causation can lead to various equilibria that involve the concentration of population at different locations.

The general case. In the previous paragraphs, I argued that the presence of the farm sector works in favor of equilibrium uniqueness, while the presence of the non-farm sector works in favor of equilibrium multiplicity. The general model of Section 3 features both sectors and as a result, intuitively, can feature either multiplicity or uniqueness. But why is the structure of this general model too complex to prevent one from theoretically characterizing when uniqueness arises? Recall that the distribution of farm population across trading places, $L_t^F(s)$, is pinned down by the equation

$$B(s)^{-\frac{1-\nu}{\nu}\frac{\epsilon(\epsilon-1)}{2\epsilon-1}} \left[\frac{L_t^F(s)}{H_t(s)}\right]^{(1-\alpha)\frac{1-\nu}{\nu}\frac{\epsilon(\epsilon-1)}{2\epsilon-1}} \bar{c}_t(s)^{\frac{(\epsilon-1)^2}{(2\epsilon-1)}} = \kappa U_t^{1-\epsilon} \cdot$$

$$\int_S \frac{B(u)^{\frac{1-\nu}{\nu}\frac{(\epsilon-1)^2}{2\epsilon-1}} \left[\frac{L_t^F(u)}{H_t(u)}\right]^{-(1-\alpha)\frac{1-\nu}{\nu}\frac{(\epsilon-1)^2}{2\epsilon-1}} \bar{c}_t(u)^{-\frac{\epsilon(\epsilon-1)}{(2\epsilon-1)}}}{(1-\nu)\left[\epsilon(1-\mu)+\mu\right]+(\epsilon-1)\mu\left[1-\nu\bar{c}_t(u)^{\beta-1}\right]} \left[\int_{\sigma_t^{-1}(u)} L_t^F(r)\,dr\right] \tau_t^N(u,s)^{1-\epsilon}\,du \tag{A.18}$$

while the distribution of farmers at locations r that trade at trading place $\sigma_t(r)$ is given by

$$L_t^F(r) = \tau_t^F(r,\sigma_t(r))^{-\frac{1}{1-\alpha}} \frac{H_t(r)}{H_t(\sigma_t(r))} \left[\frac{B(r)}{B(\sigma_t(r))}\right]^{\frac{1}{1-\alpha}} L_t^F(\sigma_t(r)) \tag{A.19}$$

where

$$\bar{c}_t(s) = \left[1 + A_t(s)^{\beta-1} B(s)^{(\beta-1)\mu} \left[\frac{L_t^F(s)}{H_t(s)}\right]^{(1-\alpha)(1-\beta)\mu}\right]^{\frac{1}{1-\beta}} \tag{A.20}$$

as a result of (3.7) and (A.9), the non-farm population of a location s is given by

$$L_t^N(s) = \frac{\nu}{(1-\nu)+\mu(\epsilon-1)\left[\mu+\epsilon(1-\mu)+\frac{\epsilon}{\bar{c}_t(s)^{1-\beta}-1}\right]^{-1}} \int_{\sigma_t^{-1}(s)} L_t^F(r)\,dr \tag{A.21}$$

[2]Allen and Arkolakis (2014) also show that the standard procedure of solving the model also becomes impossible to use when the condition guaranteeing equilibrium uniqueness does not hold. This and the presence of multiple equilibria are the reasons why I contrast the model of Section 3 with DNR, rather than with a special case of the model in which I eliminate the farm sector (and, therefore, farm hinterlands).

and the level of utility, U_t, is pinned down by national labor market clearing

$$\bar{L} = \int_S \left[L_t^F(r) + L_t^N(r)\right] dr.$$

If the distribution of farm population across trading places, $L_t^F(s)$, is unique, then the farm population distribution across other locations is uniquely determined by equation (A.19) and the non-farm population distribution is uniquely determined by equation (A.21). But does equation (A.18) pin down a unique distribution of $L_t^F(s)$? Equation (A.18) differs from equation (A.16) in three respects:

1. In the general model, people can consume the farm good. This feature of the general model is responsible for the terms $\left[\frac{L_t^F(s)}{H_t(s)}\right]^{(1-\alpha)\frac{1-\nu}{\nu}\frac{\epsilon(\epsilon-1)}{2\epsilon-1}}$ and $\left[\frac{L_t^F(u)}{H_t(u)}\right]^{-(1-\alpha)\frac{1-\nu}{\nu}\frac{(\epsilon-1)^2}{2\epsilon-1}}$ appearing on the left- and right-hand sides of equation (A.18), respectively.

2. In the general model, the farm good can be used as an intermediate in non-farm production. This feature of the general model is responsible for the term $\left[\frac{L_t^F(r)}{H_t(r)}\right]^{(1-\alpha)(1-\beta)\mu}$ appearing in equation (A.20), as well as for $\bar{c}_t(u)$ appearing in the denominator on the right-hand side of equation (A.18).

3. Farmers trade the farm good, and hence choose which trading place they want to trade at. This feature of the general model is responsible for the term $\int_{\sigma_t^{-1}(u)} L_t^F(r)\, dr$ on the right-hand side of equation (A.18).

These three additional features of the general model introduce non-linearities in equations (A.18) and (A.20), which together imply that the system is no longer a special case of equation (A.17).[3] As a result, one can no longer use the mathematical results underlying Theorem 2 in Allen and Arkolakis (2014) and characterize uniqueness of the equilibrium distribution of farm population.

Intuitively, even though the farm and non-farm sectors solely feature forces for and against equilibrium uniqueness, respectively, their complex interconnections make a clean theoretical characterization of uniqueness infeasible. Although land use provides incentives for farmers to spread out across space, they may be driven to different regions in two distinct equilibria that feature different spatial distributions of non-farm population, and therefore different spatial distributions of demand for their product. Alternatively, even though the agglomeration force works in favor of multiple equilibria in the non-farm sector, their indirect use of land through the farm input might still lead to a unique equilibrium distribution of non-farm population. In the simulations, I find the latter to be the case.

[3] With feature 1 alone, the system would still be a special case of (A.17), but not with features 2 and/or 3.

A.5 A model with home consumption

This section of the appendix presents a version of the model of Section 3 in which farmers consume goods at their residential location, not at the trading place. For non-farm goods, which farmers purchase from others, this assumption results in an extra shipping cost that they need to incur between the trading place and their residence. For the farm good, which they produce, the assumption leads to savings on shipping costs as farmers do not need to ship the fraction of the good that they consume to the trading place. In what follows, I describe the farmer's problem, as well as the set of equilibrium conditions that change relative to Section 3.

Farmers choose their production and consumption levels, their residence and their trading place to maximize utility subject to the constraints

$$q_t^F(r) = B(r)\,\ell_t^F(r)^{\alpha}\,h_t(r)^{1-\alpha}$$

and

$$\int_0^{n_t} p_t^N(s,i)\,\tau_t^N(s,r)\,x_t^N(r,s,i)\,di + R_t(r)\,h_t(r) \leq p_t^F(s)\,\tau_t^F(r,s)^{-1}\left[q_t^F(r) - x_t^F(r,s)\right] + y_t(r)$$

where the definitions of all variables are the same as in Section 3. Notice the two differences between the budget constraint presented here and the one in Section 3.1. First, the farmer needs to pay the additional cost $\tau_t^N(s,r)$ of shipping non-farm varieties home from the trading place. Second, the right-hand side has $q_t^F(r) - x_t^F(r,s)$, the difference between the quantity of the farm good produced and the quantity consumed by the farmer herself.

The farmer's indirect utility is then given by

$$\begin{aligned} U_t^F(r,s) &= \frac{p_t^F(s)\,\tau_t^F(r,s)^{-1}\,q_t^F(r)}{\left[\tau_t^N(s,r)\,P_t(s)\right]^{\nu}\left[\tau_t^F(r,s)^{-1}\,p_t^F(s)\right]^{1-\nu}} \\ &= \frac{p_t^F(s)\,\tau_t^F(r,s)^{-\nu}\,\tau_t^N(r,s)^{-\nu}\,B_t(r)\left[\frac{H_t(r)}{L_t^F(r)}\right]^{1-\alpha}}{P_t(s)^{\nu}\,p_t^F(s)^{1-\nu}}. \end{aligned} \tag{3.1'}$$

Combining this with equations (3.2), (A.3) and (A.9) implies that the trading place chosen by farmers living at location r is

$$\sigma_t(r) = argmax_{s\in S_c}\left[\frac{\tau_t^N(s,r)^{\nu}}{\tau_t^F(r,s)^{1-\nu}}\right]^{-1}\varsigma_t(r,s)^{-1}\,B(s)^{-1}\,H(s)^{-(1-\alpha)}\,L_t^F(s)^{1-\alpha} \tag{A.10'}$$

where I normalized $\tau_t^N(s,s)$ to one, and the farm population of r is

$$L_t^F(r) = \left[\frac{\tau_t^N(\sigma_t(r),r)^\nu}{\tau_t^F(r,\sigma_t(r))^{1-\nu}}\right]^{-\frac{1}{1-\alpha}} \varsigma_t(r,\sigma_t(r))^{-\frac{1}{1-\alpha}} \frac{H_t(r)}{H_t(\sigma_t(r))} \left[\frac{B(r)}{B(\sigma_t(r))}\right]^{\frac{1}{1-\alpha}} L_t^F(\sigma_t(r)). \tag{A.11'}$$

As a result, equations (A.12) and (A.13) become

$$B(s)^{-\frac{1-\nu}{\nu}\frac{\epsilon(\epsilon-1)}{2\epsilon-1}} \left[\frac{L_t^F(s)}{H_t(s)}\right]^{(1-\alpha)\frac{1-\nu}{\nu}\frac{\epsilon(\epsilon-1)}{2\epsilon-1}} \bar{c}_t(s)^{\frac{(\epsilon-1)^2}{(2\epsilon-1)}} = \kappa \bar{U}_{c(s),t}^{-\frac{\epsilon(\epsilon-1)}{2\epsilon-1}}.$$

$$\int_S \bar{U}_{c(u),t}^{-\frac{(\epsilon-1)^2}{2\epsilon-1}} \frac{B(u)^{\frac{1-\nu}{\nu}\frac{(\epsilon-1)^2}{2\epsilon-1}} \left[\frac{L_t^F(u)}{H_t(u)}\right]^{-(1-\alpha)\frac{1-\nu}{\nu}\frac{(\epsilon-1)^2}{2\epsilon-1}} \bar{c}_t(u)^{-\frac{\epsilon(\epsilon-1)}{(2\epsilon-1)}}}{(1-\nu)\left[\epsilon(1-\mu)+\mu\right]+(\epsilon-1)\mu\left[1-\nu\bar{c}_t(u)^{\beta-1}\right]}.$$

$$\left[\int_{\sigma_t^{-1}(u)} \frac{\tau_t^N(u,r)^\nu}{\tau_t^F(r,u)^{1-\nu}} L_t^F(r)\,dr\right] \tau_t^N(u,s)^{1-\epsilon}\,du \tag{A.12'}$$

and

$$L_t^N(s) = \frac{\nu}{(1-\nu)+\mu(\epsilon-1)\left[\mu+\epsilon(1-\mu)+\frac{\epsilon}{\bar{c}_t(s)^{1-\beta}-1}\right]^{-1}} \int_{\sigma_t^{-1}(s)} \frac{\tau_t^N(s,r)^\nu}{\tau_t^F(r,s)^{1-\nu}} L_t^F(r)\,dr, \tag{A.13'}$$

respectively.

A comparison of equations (A.10'), (A.11'), (A.12') and (A.13') to their counterparts (A.10), (A.11), (A.12) and (A.13) reveals how additional shipping costs $\tau_t^N(s,r)^\nu$ and shipping cost savings $\tau_t^F(r,s)^{1-\nu}$ alter the equilibrium relative to Section 3. If these shipping cost changes exactly counterbalance each other, that is, $\frac{\tau_t^N(s,r)^\nu}{\tau_t^F(r,s)^{1-\nu}} = 1$, then the equilibrium population, productivity and utility levels of the two models become identical. This is stated formally in the next proposition.

Proposition 3 (Isomorphism with model of Section 3). *Assume* $\frac{\tau_t^N(s,r)^\nu}{\tau_t^F(r,s)^{1-\nu}} = 1$ *for all* $r,s \in S$. *Then the model with home consumption is isomorphic in its evolution of population, productivity and utility to the model presented in Section 3.*

In the baseline calibration of the model of Section 3, we have $\nu = 0.75$ and $\tau_t^N(s,r) = \tau_t^F(r,s)^{\phi^N/\phi^F}$. We obtain the isomorphism whenever $\frac{\tau_t^F(r,s)^{0.75\phi^N/\phi^F}}{\varsigma_t(r,s)^{0.25}} = 1$, that is, $0.75\phi^N/\phi^F - 0.25 = 0$, or $\phi^N/\phi^F = 1/3$. The value of ϕ^N/ϕ^F used in the calibration is indeed close to this value. Therefore, Proposition 3 implies that the difference between the two models should be small, and changing the assumption about where consumption happens in space is unlikely to alter the results substantially.

A.6 Proofs

Proof of Lemma 1

The firm's optimal innovation decision (3.9) implies

$$\ell_t^I(s)^{1-\mu} = (1-\mu)^{1-\mu}(\epsilon-1)^{1-\mu} f^{1-\mu}\left[A_t(s)^{\beta-1}\left[\frac{p_t^F(s)}{w_t(s)}\right]^{(1-\beta)\mu}\bar{c}_t(s)^{\beta-1}\right]^{1-\mu}$$

from which, using (3.7), we obtain

$$\ell_t^I(s)^{1-\mu} = (1-\mu)^{1-\mu}(\epsilon-1)^{1-\mu} f^{1-\mu}\left[1-\bar{c}_t(s)^{\beta-1}\right]^{1-\mu}. \tag{A.22}$$

Equation (A.13) implies that the unit cost per wage at s can be written as

$$\bar{c}_t(s) = \left[\frac{\nu\left(1+\frac{L_t^N(s)}{\int_{\sigma_t^{-1}(s)} L_t^F(r)dr}\right)}{\left(\nu+\frac{(1-\nu)\epsilon}{\mu(\epsilon-1)}\right)\frac{L_t^N(s)}{\int_{\sigma_t^{-1}(s)} L_t^F(r)dr}-\nu\left(\frac{\epsilon}{\mu(\epsilon-1)}-1\right)}\right]^{\frac{1}{1-\beta}}.$$

Plugging this into equation (A.22) and rearranging yields the result.

Proof of Lemma 2

Equation (A.4) provides the price index at any trading place s. Combining it with equations (3.5) and (3.11) yields

$$P_t(s)^{1-\epsilon} = \left(\frac{\epsilon-1}{\epsilon}\right)^{\epsilon-1} f^{-1}\int_S \frac{\bar{c}_t(u)^{1-\epsilon}}{(\epsilon-1)\left(\mu\bar{c}_t(u)^{\beta-1}+1-\mu\right)+1} w_t(u)^{1-\epsilon} L_t^N(u)\,\tau_t^N(u,s)^{1-\epsilon}\,du. \tag{A.23}$$

Alternatively, one can express the price index at s from equation (3.2) as

$$P_t(s) = U_{c(s),t}^{-\frac{1}{\nu}} B(s)^{\frac{1-\nu}{\nu}}\left[\frac{L_t^F(s)}{H_t(s)}\right]^{-(1-\alpha)\frac{1-\nu}{\nu}} w_t(s)$$

where I used equations (A.3) and (A.9). Plugging this into the left-hand side of equation (A.23) implies

$$U_{c(s),t}^{\frac{\epsilon-1}{\nu}} B(s)^{-(\epsilon-1)\frac{1-\nu}{\nu}}\left[\frac{L_t^F(s)}{H_t(s)}\right]^{(1-\alpha)(\epsilon-1)\frac{1-\nu}{\nu}} w_t(s)^{1-\epsilon} =$$

$$\tilde{\kappa}\int_S \frac{\bar{c}_t(u)^{1-\epsilon}}{(\epsilon-1)\left(\mu\bar{c}_t(u)^{\beta-1}+1-\mu\right)+1} w_t(u)^{1-\epsilon} L_t^N(u)\,\tau_t^N(u,s)^{1-\epsilon}\,du \tag{A.24}$$

where $\tilde{\kappa} = \left(\frac{\epsilon-1}{\epsilon}\right)^{\epsilon-1} f^{-1}$.

Equation (A.6) provides the market clearing condition for each non-farm good at any trading place s. Combining it with equations (3.5) and (3.6) yields

$$\bar{c}_t(s)^{\epsilon-1} w_t(s)^{\epsilon} = \nu\epsilon^{-\epsilon}(\epsilon-1)^{\epsilon-1} f^{-1} \int_S P_t(u)^{\epsilon-1} I_t(u) \tau_t^N(s,u)^{1-\epsilon} du \tag{A.25}$$

where

$$I_t(u) = \int_{\sigma_t^{-1}(u)} p_t^F(u) \tau_t^F(r,u)^{-1} B(r) L_t^F(r)^{\alpha} H_t(r)^{1-\alpha} dr + w_t(u) L_t^N(u)$$

is the sum of farmers' and non-farm workers' income at trading place u. Equations (3.1) and (A.1) allow me to rewrite income as

$$I_t(u) = p_t^F(u) B(u) \left[\frac{H_t(u)}{L_t^F(u)}\right]^{1-\alpha} \int_{\sigma_t^{-1}(u)} L_t^F(r) dr + w_t(u) L_t^N(u)$$

from which, by equation (A.9), we obtain

$$I_t(u) = w_t(u) \left[\int_{\sigma_t^{-1}(u)} L_t^F(r) dr + L_t^N(u)\right].$$

Also, combining (A.5) with (3.10) and (3.11) yields

$$L_t^N(s) = \frac{\nu}{(1-\nu) + \mu(\epsilon-1)\left[\mu + \epsilon(1-\mu) + \frac{\epsilon}{\bar{c}_t(s)^{1-\beta}-1}\right]^{-1}} \int_{\sigma_t^{-1}(s)} L_t^F(r) dr$$

hence income can be written as

$$I_t(u) = w_t(u) \frac{\nu^{-1}\epsilon}{(\epsilon-1)\left(\mu c_t(u)^{\beta-1} + 1 - \mu\right) + 1} L_t^N(u)$$

and equation (A.25) can be written as

$$\bar{c}_t(s)^{\epsilon-1} w_t(s)^{\epsilon} = \tilde{\kappa} \int_S U_{c(u),t}^{-\frac{\epsilon-1}{\nu}} \frac{B(u)^{(\epsilon-1)\frac{1-\nu}{\nu}} \left[\frac{L_t^F(u)}{H_t(u)}\right]^{-(1-\alpha)(\epsilon-1)\frac{1-\nu}{\nu}}}{(\epsilon-1)\left(\mu\bar{c}_t(u)^{\beta-1} + 1 - \mu\right) + 1} w_t(u)^{\epsilon} L_t^N(u) \tau_t^N(s,u)^{1-\epsilon} du \tag{A.26}$$

where $\tilde{\kappa} = \left(\frac{\epsilon-1}{\epsilon}\right)^{\epsilon-1} f^{-1}$, and I used equations (3.2), (A.3) and (A.9) to substitute for the price index on the right-hand side.

In what follows, I show that equations (A.24) and (A.26) reduce to a single equation. To see this, note first that, since non-farm shipping costs are symmetric, $\tau_t^N(s,u) = \tau_t^N(u,s)$.

Second, guess that wages at location s take the form

$$w_t(s) = U_{c(s),t}^{\iota_1} B(s)^{\iota_2} \left[\frac{L_t^F(s)}{H_t(s)}\right]^{\iota_3} \bar{c}_t(s)^{\iota_4}$$

where ι_1, ι_2, ι_3 and ι_4 are constants. Plugging the guess into (A.24) and (A.26), one obtains that both of these equations hold if and only if $\iota_1 = \frac{\epsilon-1}{\nu(2\epsilon-1)}$, $\iota_2 = -\frac{1-\nu}{\nu}\frac{\epsilon-1}{2\epsilon-1}$, $\iota_3 = (1-\alpha)\frac{1-\nu}{\nu}\frac{\epsilon-1}{2\epsilon-1}$, and $\iota_4 = -\frac{\epsilon-1}{2\epsilon-1}$. Thus, wages at s can be written as

$$w_t(s) = U_{c(s),t}^{\frac{\epsilon-1}{\nu(2\epsilon-1)}} B(s)^{-\frac{1-\nu}{\nu}\frac{\epsilon-1}{2\epsilon-1}} \left[\frac{L_t^F(s)}{H_t(s)}\right]^{(1-\alpha)\frac{1-\nu}{\nu}\frac{\epsilon-1}{2\epsilon-1}} \bar{c}_t(s)^{-\frac{\epsilon-1}{2\epsilon-1}}$$

and equations (A.24) and (A.26) reduce to equation (A.12).

Proof of Proposition 1

Rearranging equation (A.13) yields location s's effective hinterland size as

$$\tilde{H}_t(s) = \frac{\int_{\sigma_t^{-1}(s)} L_t^F(r)\,dr}{L_t^N(s)} = \frac{1}{\nu}\left[(1-\nu) + \mu(\epsilon-1)\left(\mu + \epsilon(1-\mu) + \frac{\epsilon}{\bar{c}_t(s)^{1-\beta} - 1}\right)^{-1}\right]$$

from which, by rearranging,

$$\bar{c}_t(s)^{1-\beta} - 1 = \frac{\epsilon}{\frac{\mu(\epsilon-1)}{\nu\tilde{H}_t(s)-(1-\nu)} - \mu - \epsilon(1-\mu)}. \tag{A.27}$$

To show part i of the proposition, combine equation (A.27) with (3.7) and (A.9) to get

$$L_t^F(s) = \left[\frac{\epsilon}{\frac{\mu(\epsilon-1)}{\nu\tilde{H}_t(s)-(1-\nu)} - \mu - \epsilon(1-\mu)}\right]^{\frac{1}{(1-\alpha)(1-\beta)\mu}} A_t(s)^{\frac{1}{(1-\alpha)\mu}} B(s)^{\frac{1}{1-\alpha}} H_t(s) \tag{A.28}$$

which implies that the farm population of s is increasing in $\tilde{H}_t(s)$ as

$$\begin{aligned}\frac{dL_t^F(s)}{d\tilde{H}_t(s)} =& \frac{1}{(1-\alpha)(1-\beta)\mu}\left[\frac{\epsilon}{\frac{\mu(\epsilon-1)}{\nu\tilde{H}_t(s)-(1-\nu)} - \mu - \epsilon(1-\mu)}\right]^{\frac{1}{(1-\alpha)(1-\beta)\mu}-1} A_t(s)^{\frac{1}{(1-\alpha)\mu}} \cdot \\ & B(s)^{\frac{1}{1-\alpha}} H_t(s) \frac{\epsilon\frac{\mu(\epsilon-1)}{\left(\nu\tilde{H}_t(s)-(1-\nu)\right)^2}\nu}{\left(\frac{\mu(\epsilon-1)}{\nu\tilde{H}_t(s)-(1-\nu)} - \mu - \epsilon(1-\mu)\right)^2} > 0.\end{aligned}$$

The farm population of any other location r in s's hinterland is also increasing in $\tilde{H}_t(s)$

since, by equation (A.11),

$$\frac{dL_t^F(r)}{d\tilde{H}_t(s)} = \tau_t^F(r,s)^{-\frac{1}{1-\alpha}} \frac{H_t(r)}{H_t(s)} \left[\frac{B(r)}{B(s)}\right]^{\frac{1}{1-\alpha}} \frac{dL_t^F(s)}{d\tilde{H}_t(s)} > 0.$$

To show part ii of the proposition, use the definition of $\tilde{H}_t(s)$ to obtain

$$L_t^N(s) = \tilde{H}_t(s)^{-1} \int_{\sigma_t^{-1}(s)} L_t^F(r)\, dr$$

then use (A.11) and (A.28) to get

$$\begin{aligned} L_t^N(s) =& \tilde{H}_t(s)^{-1} \left[\frac{\epsilon}{\frac{\mu(\epsilon-1)}{\nu\tilde{H}_t(s)-(1-\nu)} - \mu - \epsilon(1-\mu)}\right]^{\frac{1}{(1-\alpha)(1-\beta)\mu}} A_t(s)^{\frac{1}{(1-\alpha)\mu}} \cdot \\ & \int_{\sigma_t^{-1}(s)} \tau_t^F(r,s)^{-\frac{1}{1-\alpha}} B(r)^{\frac{1}{1-\alpha}} H_t(r)\, dr \end{aligned}$$

from which

$$\begin{aligned} L_t^N(s) = & \left[\frac{\epsilon}{\frac{\mu(\epsilon-1)}{\nu\tilde{H}_t(s)^{1-(1-\alpha)(1-\beta)\mu}-(1-\nu)\tilde{H}_t(s)^{-(1-\alpha)(1-\beta)\mu}} - (\mu+\epsilon(1-\mu))\tilde{H}_t(s)^{(1-\alpha)(1-\beta)\mu}}\right]^{\frac{1}{(1-\alpha)(1-\beta)\mu}} \cdot \\ & A_t(s)^{\frac{1}{(1-\alpha)\mu}} \int_{\sigma_t^{-1}(s)} \tau_t^F(r,s)^{-\frac{1}{1-\alpha}} B(r)^{\frac{1}{1-\alpha}} H_t(r)\, dr \end{aligned} \tag{A.29}$$

which implies $\frac{dL_t^N(s)}{d\tilde{H}_t(s)} > 0$ since $0 < (1-\alpha)(1-\beta)\mu < 1$. Thus, the non-farm population of s is increasing in $\tilde{H}_t(s)$.

To show part iii of the proposition, apply the definition of $\tilde{H}_t(s)$ in equation (3.21) to get

$$\ell_t^I(s)^{1-\mu} = \rho\left[\left(1+\tilde{H}_t(s)^{-1}\right)^{-1} - (1-\nu)\right]^{1-\mu} \tag{A.30}$$

which implies

$$\frac{d\ell_t^I(s)^{1-\mu}}{d\tilde{H}_t(s)} = (1-\mu)\rho^{\frac{1}{1-\mu}}\ell_t^I(s)^{-\mu} \frac{\tilde{H}_t(s)^{-2}}{\left(1+\tilde{H}_t(s)^{-1}\right)^2} > 0.$$

That is, innovation at s is increasing in $\tilde{H}_t(s)$ as well.

Proof of Proposition 2

In the special case of the model presented in Section A.3, equations (3.7) and (A.9) together imply

$$L_t^F(s) = \left[\bar{c}_t(s)^{1-\beta} - 1\right]^{\frac{1}{(1-\alpha)(1-\beta)\mu}} A_t(s)^{\frac{1}{(1-\alpha)\mu}} B(s)^{\frac{1}{1-\alpha}} \tau_t(s)^{-\frac{1}{1-\alpha}} H_t(s) \tag{A.31}$$

while equation (A.12) simplifies to

$$\bar{c}_t(s)^{\frac{(\epsilon-1)^2}{2\epsilon-1}} = \kappa \bar{U}_t^{1-\epsilon}\left(1 + b_t(s)^{1-\epsilon}\right) \frac{\bar{c}_t(s)^{-\frac{\epsilon(\epsilon-1)}{2\epsilon-1}}}{(\epsilon-1)\mu\left[1 - \bar{c}_t(s)^{\beta-1}\right]} S_t(s) L_t^F(s). \tag{A.32}$$

Plugging (A.31) into (A.32), rearranging and taking logs yields

$$\begin{aligned}(\epsilon-1)\log\bar{c}_t(s) =& \log\left(\frac{\kappa}{(\epsilon-1)\mu}\right) + (1-\epsilon)\bar{U}_t + \log\left(1 + b_t(s)^{1-\epsilon}\right) - \log\left[1 - \bar{c}_t(s)^{\beta-1}\right] \\ &+ \frac{1}{(1-\alpha)(1-\beta)\mu}\log\left[\bar{c}_t(s)^{1-\beta} - 1\right] + \log S_t(s) + \frac{1}{(1-\alpha)\mu}\log A_t(s) \\ &+ \frac{1}{1-\alpha}\log B(s) - \frac{1}{1-\alpha}\log\tau_t(s) + \log H_t(s).\end{aligned} \tag{A.33}$$

To show part 1 of the proposition, take the derivative of (A.33) with respect to $\log\tau_t(s)$:

$$(\epsilon-1)\frac{d\log\bar{c}_t(s)}{d\log\tau_t(s)} = -\frac{(1-\beta)\bar{c}_t(s)^{\beta-1}}{1-\bar{c}_t(s)^{\beta-1}}\frac{d\log\bar{c}_t(s)}{d\log\tau_t(s)} + \frac{1}{(1-\alpha)\mu}\frac{\bar{c}_t(s)^{1-\beta}}{\bar{c}_t(s)^{1-\beta}-1}\frac{d\log\bar{c}_t(s)}{d\log\tau_t(s)} - \frac{1}{1-\alpha}$$

Rearranging yields

$$\frac{d\log\bar{c}_t(s)}{d\log\tau_t(s)} = -\frac{1}{(1-\alpha)\left(\epsilon - 1 - \frac{1}{(1-\alpha)\mu} - \frac{\frac{1}{(1-\alpha)\mu}-(1-\beta)}{\bar{c}_t(s)^{1-\beta}-1}\right)} < 0$$

where the inequality comes from the assumption $\bar{c}_t(s)^{1-\beta} - 1 > \frac{\frac{1}{(1-\alpha)\mu}-(1-\beta)}{\epsilon-1-\frac{1}{(1-\alpha)\mu}}$. Taking the derivative of (A.27) with respect to $\log\bar{c}_t(s)$ and rearranging, we also obtain

$$\frac{d\log\tilde{H}_t(s)}{d\log\bar{c}_t(s)} = \frac{(1-\beta)\bar{c}_t(s)^{1-\beta}\left(\mu(\epsilon-1)\tilde{H}_t(s)^{-1} - \mu - \epsilon(1-\mu)\right)^2}{\mu\epsilon(\epsilon-1)\tilde{H}_t(s)^{-1}} > 0.$$

Pulling the two derivatives together, we get, by the chain rule,

$$\frac{d\log\tilde{H}_t(s)}{d\log\tau_t(s)} = \frac{d\log\tilde{H}_t(s)}{d\log\bar{c}_t(s)}\frac{d\log\bar{c}_t(s)}{d\log\tau_t(s)} < 0$$

which proves that the size of the effective hinterland is decreasing in $\tau_t(s)$. This is part 1 of the proposition.

The proof of parts 2 and 3 is analogous, with the exception that one needs to take the derivative of (A.33) with respect to $\log S_t(s)$ and $\log b_t(s)$, respectively.

Proof of Corollary of Proposition 1

In the case of international trade costs, the result is an immediate consequence of Proposition 1 due to equations (A.28), (A.29) and (A.30).

In the case of railroad construction, we also need to take into account the direct effect it has on $\tau_t^F(r,s)$ in equations (A.28), (A.11) and (A.29). However, since railroads lower all $\tau_t^F(r,s)$, this direct effect amplifies the indirect effect that railroads have through increasing effective hinterland size. Hence, $L_t^F(r)$, $L_t^N(s)$ and $\ell_t^I(s)$ all increase.

In the case of the expansion of political borders, there might be a direct effect on the set of locations in s's hinterland, $\sigma_t^{-1}(s)$, and thus on $L_t^N(s)$. But since $\sigma_t^{-1}(s)$ cannot shrink by the fact that political borders expand – as well as by equation (A.10) –, we again get a direct effect that amplifies the indirect effect that border expansion has through increasing effective hinterland size. Hence, $L_t^F(r)$, $L_t^N(s)$ and $\ell_t^I(s)$ all increase in that case as well.

B Data appendix

This appendix describes the datasets used to document the major patterns of pre-Civil War U.S. urban history, to take the model to the data, and to evaluate the model's fit. I use geographic data on the location of the sea, navigable rivers, canals and lakes, as well as railroads to calculate shipping costs and to quantify the importance of trading routes in city location. I use census data on county, city and town locations and populations to calibrate the model and to evaluate how well the model fits the evolution of population seen in the data.

My unit of observation is a *cell* in a 20 by 20 arc minute grid of the United States. I create this grid of the U.S. using Geographic Information Software (GIS), and combine it with other sources of geographic data to determine whether any given cell is at a waterway or at a railroad, and to calculate the agricultural productivity, natural amenities, and population of the cell. In what follows, I provide additional details on this procedure.

Waterways. I use the ESRI Map of U.S. Major Waters and the 20 by 20 arc minute grid of the U.S. to determine whether a grid cell is at the sea. In particular, I regard a cell as being at the sea if a positive fraction of its area is in the sea.

I follow Donaldson and Hornbeck's (2016) definition of navigable rivers, lakes and canals, who, in turn, borrow the definition from Fogel (1964). Combining the definition with the

ESRI Map of U.S. Major Waters and the 20 by 20 arc minute grid of the U.S., I classify each cell based on whether it contains a navigable waterway. As canals were gradually constructed during the 19th century, I do this classification of cells separately for every time period t, using the set of canals that were already open at t. Table 10 provides a list of navigable canals, along with their locations and opening dates.

Railroads. The website *http://oldrailhistory.com* includes maps of the U.S. railroad network in 1835, 1840, 1845, 1850 and 1860.[4] I georeference these maps to the 20 by 20 arc minute grid of the U.S., and classify each cell depending on whether it contained some railroads in any given period t between 1835 and 1860.[5]

Agricultural productivity. I collect high-resolution data on agricultural yields from the Food and Agriculture Organization's Global Agro-Ecological Zones database (FAO GAEZ). This database contains the potential yield of various crops at a 5 by 5 arc minute spatial resolution, under different irrigation and input conditions. These yields come in the form of an index whose value, in its raw form, ranges between 0 and 10,000 for each crop-location pair. This makes the yields comparable across crops and across locations. To provide the best possible approximation to 19th-century productivity, I calculate the yields under the assumption of no irrigation and low input levels. I use data on the potential yields of cereals, cotton, sugar cane, sweet potato, tobacco, and white potato.[6] I aggregate the data to the 20 by 20 arc minute level by calculating the average productivity of each crop within each 20 by 20 minute cell. Table 11 provides summary statistics of productivity for each crop.

Natural amenities. The FAO GAEZ dataset also includes data on natural amenities as they heavily influence agricultural yields. I select five climate variables that are the closest to standard measures of natural amenities in the literature (see, for instance, Desmet and Rossi-Hansberg, 2013): the mean annual temperature, the annual temperature range, the number of days with minimum temperature below 5 °C, the number of days with mean temperature above 10 °C, and the annual precipitation.[7] To be as close as possible to 19th-century conditions, I use the earliest data available (1961 to 1990 for the annual temperature range, and 1960 for the other variables). I aggregate the variables to the 20

[4] Although the first railroads started to be built in the late 1820s, there only existed a small number of short and disconnected segments in 1830. Therefore, it is reasonable to assume that no U.S. location had access to a rail network until 1830.

[5] Lacking the map of the network in 1855, I need to approximate it by the network in 1850.

[6] According to the 1860 Census of Agriculture, these were the six major crops grown in the United States.

[7] Although these variables are close to the ones considered in the literature, they do not exactly coincide with them. Therefore, as a robustness check, I collect the climate variables used in Desmet and Rossi-Hansberg (2013) from *weatherbase.com* for a 845-element subset of U.S. grid cells. Using these alternative variables does not alter the results substantially. These results are available from the author upon request.

Table 10: Navigable canals constructed between 1790 and 1860

Year of opening	Canal name	Canal was constructed to connect...
1823	Champlain Canal	Lake Champlain and Hudson River
1825	Erie Canal	Lake Erie and Hudson River
1827	Schuylkill Canal	Port Carbon, PA and Philadelphia
1828	Erie Canal, Oswego branch	Erie Canal and Lake Ontario
1828	Union Canal	Middletown, PA and Reading, PA
1828	Delaware and Hudson Canal	Delaware River and Hudson River
1828	Farmington Canal	New Haven, CT and interior of Connecticut
1828	Blackstone Canal	Worcester, MA and Providence, RI
1829	Chesapeake and Delaware Canal	Chesapeake Bay and Delaware River
1831	Morris Canal	Phillipsburg, NJ and Jersey City, NJ
1832	Ohio and Erie Canal	Lake Erie and Ohio River
1832	Pennsylvania Canal System, Delaware Division	Easton, PA and Bristol, PA
1832	Cumberland and Oxford Canal	lakes in Southern Maine and Portland, MA
1834	Delaware and Raritan Canal	Delaware River and New Brunswick, NJ
1834	Chenango Canal	Binghamton, NY and Utica, NY
1835	Pennsylvania Canal System	several rivers and canals in Pennsylvania
1840	Susquehanna and Tidewater Canal	Wrightsville, PA and Chesapeake Bay
1840	Pennsylvania and Ohio Canal	Ohio and Erie Canal and Beaver and Erie Canal
1840	James River and Kanawha Canal	Lynchburg, VA and Richmond, VA
1841	Genesee Valley Canal	Dansville, NY and Rochester, NY
1844	Beaver and Erie Canal	Lake Erie and Ohio River
1845	Miami and Erie Canal	Lake Erie and Cincinatti, OH
1847	Whitewater Canal	Ohio River and Lawrenceburg, IN
1848	Illinois and Michigan Canal	Lake Michigan and Illinois River
1848	Wabash and Erie Canal, section 1	Miami and Erie Canal and Terre Haute, OH
1848	Sandy and Beaver Canal	Ohio and Erie Canal and Ohio River
1850	Chesapeake and Ohio Canal	Cumberland, MD and Washington, D.C.
1853	Wabash and Erie Canal, section 2	Ohio River and Terre Haute, OH
1855	Black River Canal	Black River and Erie Canal
1858	Chemung and Junction Canals	Erie Canal and Pennsylvania Canal

Table 11: Productivity of the six main U.S. crops

Name of crop	Minimum	Maximum	Mean	Standard deviation
Cereals	0	10,000	4,332	2,326
Cotton	0	9,549	1,524	2,150
Sugar cane	0	9,966	266	937
Sweet potato	0	10,000	677	1,529
Tobacco	0	9,703	2,376	2,257
White potato	0	9,596	2,979	1,897

The four columns of this table show the minimum, the maximum, the unweighted mean and the unweighted standard deviation of grid cell-level yields for the six major crops in the United States. Filters "low-input level" and "rain-fed" have been applied for each crop. Source: FAO GAEZ.

by 20 arc minute level using the same procedure as the one used for productivity. Table 12 provides summary statistics of the natural amenity variables.

County, city and town populations. The National Historical Geographic Information System (NHGIS) provides census data on county populations for 1790, 1800, 1810, 1820, 1830, 1840, 1850 and 1860, along with maps of county boundaries.[8] I use the county population data to calculate the population of each 20 by 20 arc minute grid cell. For each census year, I transform the county map into a raster of 2 by 2 arc minute cells, and allocate the population of each county equally across the small cells it occupies. Next, I calculate the population of each 20 by 20 minute cell by summing the population levels of 2 by 2 minute cells inside the cell. Finally, I obtain city and town populations from a census database that provides the population of settlements above 2,500 inhabitants in each census year,[9] while I use Google Maps to determine the geographic location of each town and city.

Large regions. Based on the boundaries of U.S. states today, I assign each 20 by 20 arc minute grid cell to the state to which its centroid belongs. Next, I assign the cell to one of the four large U.S. regions: the Northeast, the South, the Midwest or the West, following the mapping of states to regions in Caselli and Coleman (2001). Therefore, the Northeast constitutes of Connecticut, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island and Vermont; the South includes Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia and West Virginia; the Midwest constitutes of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota and Wisconsin; and the remaining states

[8]The database is available at *nhgis.org*. Source: Minnesota Population Center. *National Historical Geographic Information System: Version 2.0.* Minneapolis, MN: University of Minnesota 2011.

[9]This database is available at *census.gov/population/www/documentation/twps0027/twps0027.html.*

Table 12: Natural amenity variables

Variable	Minimum	Maximum	Mean	Standard deviation
Mean annual temperature (°C)	-2.1	24.4	10.9	5.1
Annual temperature range (°C)	5.3	37.4	24.6	5.4
Number of days with minimum temperature below 5 °C	0	251	111.2	61.2
Number of days with mean temperature above 10 °C	23.4	365	202.4	58.4
Annual precipitation (mm)	44.8	2634.8	695.2	417.4

The four columns of this table show the minimum, the maximum, the unweighted mean and the unweighted standard deviation of grid-cell level natural amenity variables for the United States. All data are for 1960, except annual temperature range which is for the period between 1961 and 1990. Source: FAO GAEZ.

belong to the West.

U.S. land. I also use the NHGIS database to calculate the fraction of cells that is covered by land *and* is part of the U.S. in any census year. In particular, I determine whether each 2 by 2 arc minute cell was part of U.S. territory in census year t. Then I calculate the land area of each 20 by 20 minute cell as the fraction of 2 by 2 minute cells inside the 20 by 20 minute cell that were part of U.S. territory at t.

For periods between census years (1795, 1805, 1815, 1825, 1835, 1845 and 1855), I use the fact that no significant border change took place between 1790 and 1800, between 1805 and 1815, between 1825 and 1840, and between 1855 and 1860. Therefore, I can use the 1790 (or the 1800) distribution of land in 1795, the 1810 distribution in 1805 and 1815, the 1830 (or 1840) distribution in 1825 and 1835, and the 1860 distribution in 1855. This leaves me with the task of obtaining the distribution in 1845. To accomplish this, I georeference a map showing 1845 borders to the 20 by 20 minute grid, and determine whether the centroid of each grid cell was in U.S. territory in 1845.

C Slow migration to the West

C.1 Setup

The slow settlement of the West could be another reason for the emergence of cities at trading routes. To see if such a slow migration process could quantitatively generate the disproportionate emergence of cities at trading routes, I consider a simple process in which a resident of location j moves to location i next period with probability $\pi_{ji} \in [0,1]$. That is, I assume that the population distribution evolves according to the equation

$$\ell_{i,t+1} = \nu_{t+1} \sum_{S_t} \ell_{jt} \pi_{ji}$$

where i and j index locations, S_t denotes the set of U.S. locations in period t, and $\ell_{i,t}$ is the population (per unit of land) of location i at time t. ν_{t+1} is a number that guarantees that population levels sum to total population in period $t+1$.

To generate slow transitions, I assume that the probability of moving from j to i is negatively related to the moving cost between these two locations:

$$\pi_{ji} = e^{-m_{ji}}$$

where m_{ji} is the cost of moving from j to i. Moving from j to i requires passing through a set of locations that connect i to j. I denote the cost of passing through a specific location k by $\xi_k > 0$. This means that the total cost of moving from j to i is

$$m_{ji} = \sum_{g_{ji}} \xi_k$$

where g_{ji} denotes the least-cost route from j to i. Note that this formulation implies that staying put is costless but moving to another location is costly. As a result, people mostly stay where they have lived in the past and occupy previously unoccupied Western regions at a slow pace. The speed of the transition is driven by the magnitude of the moving cost parameters ξ_k.

I simulate this benchmark process of slow migration to see if it is able to replicate the disproportionately frequent emergence of population clusters at water, confluences or railroads. Simulation of the process requires choosing the length of a time period, defining the set of locations for each period, choosing an initial (period 1) distribution of population across locations, and calibrating the values of moving cost parameters ξ_k. I choose the length of a period to be five years.[10] I define the set of locations in period t as the set of 20 by 20 arc minute grid cells that were part of the U.S. in period t.[11] To obtain the initial distribution of population, I assign the population of each county from the 1790 census to the grid cells it occupies, based on the share of land belonging to each cell.

For the key parameters ξ_k, I consider two calibrations. In one calibration, I make the simplest possible assumption and set ξ_k equal across locations: $\xi_k = \xi$. This implies that the cost of moving between two locations is proportional to the geographic distance between them. As the magnitude of ξ drives the speed of transition, I choose the value of ξ such that the mean center of population moves as much to the West in the model as in the data. This implies $\xi = 0.016$. Under this value of moving costs, the probability that a person moves 50 miles from her current residence next period is 45% of the probability that she does not move.

[10]Using a different period length does not alter the results substantially.

[11]That is, I incorporate political border changes between 1790 and 1860 by allowing the set of inhabitable locations to change across periods.

One concern about the assumption of equal moving costs is that migration may have been, similarly to trade, less costly along water routes than inland. I address this issue in the second calibration by making moving costs heterogeneous across locations. In particular, I assume that ξ_k equals ξ^W if location k is located at a waterway, and ξ^I otherwise. Since I have two structural parameters now, I need to calibrate them to two moments in the data. To find the right data moments, note that the speed of the transition now depends on the *average* of these costs. On the other hand, the *difference* between inland and water costs drives the extent to which people locate near rivers, hence the concentration of population in space. Therefore, I calibrate average costs $\frac{\xi^W+\xi^I}{2}$ to match the movement in the mean center of population by 1860, while I calibrate the difference in costs $\xi^I - \xi^W$ to match the concentration of population, measured by Theil index, in 1860. The resulting estimates are $\xi^W = 0.01575$ and $\xi^I = 0.01725$. This implies that moving 50 miles inland decreases the probability of moving (and, therefore, the number of migrants) by 6 percent more than moving 50 miles along water.[12]

Given that cities appeared in 75 grid cells in the data between 1790 and 1860, I define cities in the simulated data as the 75 grid cells with largest population in 1860. Then I calculate the fraction of simulated cities at water, confluences and railroads. Next, I control for the amenities and productivity of each location and estimate equation (2.1) on the simulated data.

C.2 Simulation results

Table 13 presents the results and confirms that the process of slow migration is unable to replicate the disproportionately high share of cities near trading routes. In particular, the top panel of Table 13 shows that the fraction of cities forming near a waterway, a confluence or a railroad is substantially lower under both calibrations than in the data.

The bottom panel presents the results of estimating equation (2.1) on the simulated data (columns 4 to 9), and contrasts them with the estimates coming from the actual data (columns 1 to 3). As columns (1) to (3) show, the probability that a city appears at water, confluences or railroads conditional on amenities and productivity is significantly higher in the data than the probability that it appears elsewhere. This is never the case in the simulated data, no matter whether moving costs are only a function of geographic distance (columns 4 to 6), or are lower along water routes than inland (columns 7 to 9).

To sum up, the slow migration process presented in this appendix is unable to replicate the disproportionate emergence of cities near good trading opportunities. To match these facts in the data, it seems essential to have a framework in which incentives to trade attract people to these locations.

[12] To find the least-cost routes in an efficient way, I apply the Fast Marching Algorithm for any parameter vector (ξ^W, ξ^I).

Table 13: Slow migration process: simulation results

		Slow migration process	
Fraction of cities at...	Data	Equal cost calibration	Heterogeneous cost calibration
water	98.6%	49.3%	62.7%
confluence	85.5%	33.3%	48.0%
railroad	95.7%	60.0%	58.7%

Dependent variable: newcity

	Data			Equal cost calibration			Heterogeneous cost calibration		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
water	0.018**			-0.006			0.003		
	(0.002)			(0.004)			(0.003)		
confluence		0.029**			-0.004			0.008	
		(0.004)			(0.004)			(0.004)	
railroad			0.030**			0.008			0.007
			(0.004)			(0.004)			(0.004)
Prod & amenities	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	7641	7641	7641	7641	7641	7641	7641	7641	7641

The first column of the top table shows the fraction of 1860 U.S. cities that formed in 20 by 20 arc minute grid cells located at water, at a confluence, or at a railroad in the data (also reported in Table 1). The second and third columns show the same numbers in the slow migration model of Appendix C. "Equal cost" refers to the calibration in which moving costs are equal across locations; "heterogeneous cost" refers to the calibration in which moving costs are lower along water than inland. The bottom table presents the results of estimating equation (2.1) on actual data as well as on simulated data coming from the slow migration model. The unit of observation is a 20 by 20 arc minute grid cell in the 1860 U.S. The dependent variable, *newcity*, is a dummy variable equal to 1 if a city formed in the cell by 1860 in the data (columns 1 to 3), or in the simulation (columns 4 to 9). The three independent variables, *water*, *confluence* and *railroad*, are dummy variables equal to 1 if the cell is located at water, at a confluence, or at a railroad, respectively. "Prod & amenities" refer to productivity and amenity controls. A cell is located at water if the cell, or the cell next to it, has a navigable river, canal, lake, or the sea in 1860. It is located at a confluence if the cell, or a cell next to it, is surrounded by at least 3 cells with water in 1860. It is located at a railroad if the cell, or a cell next to it, was part of the railroad network in 1860. Heteroskedasticity-robust standard errors in parentheses. *: significant at 5%; **: significant at 1%. Source: U.S. census, ESRI Map of U.S. Major Waters, *oldrailhistory.com* and FAO GAEZ.

D The DNR model

In this appendix, I develop an extension of the DNR model that I can use as a benchmark to which I compare the model of Section 3. Relative to Desmet, Nagy and Rossi-Hansberg (2018), I make three necessary but minor modifications that make the DNR model comparable to the model of Section 3. First, I allow the exogenous supply of land to change over time, so that I can incorporate changes in U.S. political borders.[13] Second, I allow total world population to grow as population growth is a prominent feature of the data. Finally, I allow shipping costs across locations to change over time as the U.S. transportation network develops. Appendix D.1 provides a detailed description of this model.

To take the DNR model to the data, I apply a procedure that follows, as closely as possible, the calibration strategy developed in Desmet et al. (2018). This calibration strategy involves matching the initial (1790) population distribution to recover location-specific productivity levels. It also involves matching the changes in population between 1790 and the next observed year (1800) to back out the levels of moving costs across locations. Intuitively, a location toward which people move is, everything else fixed, identified as a location with low costs of entry. Importantly, the calibration leaves the evolution of the population distribution after 1800 untargeted. As a result, I can simulate the DNR model until 1860, and compare how well it can replicate the evolution of the population distribution and the locations of cities to the model of Section 3. Appendix D.2 describes the procedure of taking the DNR model of the data. Appendix D.3 provides details on the simulation, while Appendix D.4 discusses the simulation results. Appendix D.5 provides robustness.

D.1 Setup

The setup is identical to the one in Desmet et al. (2018), with three necessary but minor modifications that make the model comparable to the model of Section 3. First, I allow the exogenous supply of land to change over time, so that I can incorporate changes in U.S. political borders that made land in the West available. Hence, I assume that the world is a subset S of a two-dimensional surface, and the supply of land at location $r \in S$ is given by $H_t(r)$, potentially changing across periods $t \in \{0, 1, ...\}$. Second, I allow total world population to grow as population growth is a prominent feature of the data. Hence, an exogenous number of $\bar{L}_t$ agents occupy the world in period t, and each of them supplies one unit of labor. Third, I allow shipping costs across locations to change over time as the transportation network develops.

The rest of the model is the same as Desmet et al. (2018). Most importantly, and unlike in the model of Section 3, the economy in DNR is assumed to consist of one sector that produces a continuum of differentiated varieties $\omega \in [0, 1]$.

[13]Desmet et al. (2021) also consider an extension of the DNR framework with time-varying land supply.

D.1.1 Consumption and location choice

Each agent derives utility from consuming the varieties produced by the single sector, from consuming amenities at her location and from her idiosyncratic taste for the location. Her utility is also affected by the costs of moving across locations in the past. As a result, the utility of agent i living at r in period t and having a history of past locations $\{r_0, ..., r_{t-1}\}$ is given by

$$u_t^i (r_0, ..., r_{t-1}, r) = a_t (r) \left[\int_0^1 c_t^\omega (r)^\rho \, d\omega \right]^{\frac{1}{\rho}} \varepsilon_t^i (r) \prod_{s=1}^{t} m (r_{s-1}, r_s)^{-1}$$

where $a_t (r)$ is the level of amenities at r, $c_t^\omega (r)$ is the agent's consumption of variety ω, ρ is a measure of substitutability across varieties (such that $1/(1-\rho)$ is the elasticity of substitution), $\varepsilon_t^i (r)$ is the agent's taste for location r, and $m (r_{s-1}, r_s)$ is the moving cost from location r_{s-1} to location r_s. Local amenities are given by $a_t (r) = \bar{a} (r) \bar{L}_t (r)^{-\lambda}$, where $\bar{a} (r)$ is the time-invariant fundamental amenity of the location, while $\bar{L}_t (r)^{-\lambda}$ is a congestion externality that makes locations with higher population density, $\bar{L}_t (r)$, less attractive.

For tractability, moving costs need to be modeled as a product of an origin-specific term and a destination-specific term, that is,

$$m (r, s) = m_1 (r) \, m_2 (s) \, .$$

It is reasonable to assume that staying put is costless, implying $m (r, r) = 1$ for any r and hence

$$m_1 (r) = m_2 (r)^{-1} \, .$$

As a result, the spatial distribution of moving costs is fully characterized by the function $m_2 (\cdot)$.

Location tastes $\varepsilon_t^i (r)$ are drawn independently across agents, locations and time from a Fréchet distribution with shape parameter $1/\Omega$. This implies that the equilibrium number of people choosing to reside at location r is given by

$$H_t (r) \, \bar{L}_t (r) = \frac{u_t (r)^{1/\Omega} \, m_2 (r)^{-1/\Omega}}{\int_S u_t (s)^{1/\Omega} \, m_2 (s)^{-1/\Omega}} \bar{L}_t \tag{D.1}$$

where $u_t (r)$ is the agent's utility stemming from her consumption of local amenities and varieties, thus equal to

$$u_t (r) = a_t (r) \, \frac{w_t (r) + R_t (r) / \bar{L}_t (r)}{P_t (r)} \tag{D.2}$$

where $w_t(r)$ is the agent's labor income, $R_t(r)$ denotes local land rents (per unit of land) redistributed to agents at r, and $P_t(r)$ denotes the CES price index of varieties at r.

D.1.2 Technology

At any location r, there is a large number of competitive firms producing each variety ω. Each of these firms has access to the production function

$$q_t^{\omega}(r) = \phi_t^{\omega}(r)^{\gamma_1} z_t^{\omega}(r) L_t^{\omega}(r)^{\mu}$$

where $q_t^{\omega}(r)$ denotes the firm's output per unit of land, $L_t^{\omega}(r)$ denotes labor per unit of land used by the firm in the production of ω, $z_t^{\omega}(r)$ is a location-, variety- and time-specific productivity shifter, and $\phi_t^{\omega}(r)$ is the amount of innovation conducted by the firm to increase its productivity. Innovation requires labor, such that the firm needs to hire $\nu_t \phi_t^{\omega}(r)^{\xi}$ workers to innovate $\phi_t^{\omega}(r)$ in period t. Idiosyncratic productivity shifters $z_t^{\omega}(r)$ are drawn independently across locations, varieties and time from a Fréchet distribution with cdf

$$Pr\left[z_t^{\omega}(r) \leq z\right] = e^{-T_t(r) z^{-\theta}}$$

such that the scale parameter of the distribution is given by $T_t(r) = \tau_t(r) \bar{L}_t(r)^{\alpha}$. $\bar{L}_t(r)^{\alpha}$ is an agglomeration externality that makes the location's productivity increase in population density, while $\tau_t(r)$ evolves according to the law of motion

$$\tau_t(r) = \phi_{t-1}(r)^{\theta\gamma_1} \left[\int_S \eta \tau_{t-1}(s)\, ds\right]^{1-\gamma_2} \tau_{t-1}(r)^{\gamma_2} \tag{D.3}$$

where the term $\left[\int_S \eta \tau_{t-1}(s)\, ds\right]^{1-\gamma_2}$ allows for global diffusion of technology as long as $\gamma_2 < 1$, while $\phi_{t-1}(r)$ stands for average innovation conducted by local firms at time $t-1$.[14] As a result, innovation at a location not only increases the location's current productivity but also shifts up the distribution of local productivity draws next period. Thus, innovation can increase not only short-run productivity but also long-run growth. At the same time, just like in the model of Section 3, firms in the DNR model are aware that their innovation spills over to their local competitors next period, driving down their future profits to zero. Hence, they choose a level of innovation that maximizes their current-period profits. In other words, firms solve a static optimization problem each period, which makes the model computationally tractable even though it features dynamics with a large number of asymmetric locations. The optimal level of innovation chosen by each firm at location r

[14] In equilibrium, all firms in a given location conduct the same level of innovation. Hence, it does not matter which moment of the distribution of innovation levels I consider; for completeness, I assume that it is the average level of innovation that matters for the evolution of productivity.

is given by

$$\phi_t(r) = \left[\frac{\gamma_1}{\nu_t(\mu\xi + \gamma_1)}\bar{L}_t(r)\right]^{1/\xi}. \tag{D.4}$$

Varieties are tradable across locations, but trade is subject to shipping costs. Shipping costs take the iceberg form. If q units of a variety are shipped from location r to location s, the quantity that actually arrives at s is given by $\frac{q}{\varsigma_t(r,s)}$. Thus, $\varsigma_t(r,s) \geq 1$ denotes the iceberg shipping cost between r and s.

D.1.3 Equilibrium

Competition among firms implies that the price of any variety produced at r and sold at s becomes equal to its marginal cost:

$$p_t^{\omega}(r,s) = \frac{mc_t(r)\,\varsigma_t(r,s)}{z_t^{\omega}(r)}$$

where

$$mc_t(r) = \mu^{-\mu}\left(\frac{\nu_t\xi}{\gamma_1}\right)^{1-\mu}\left[\frac{\gamma_1 R_t(r)}{w_t(r)\,\nu_t(\xi(1-\mu)-\gamma_1)}\right]^{(1-\mu)-\frac{\gamma_1}{\xi}}\frac{w_t(r)\,\varsigma_t(r,s)}{z_t^{\omega}(r)}. \tag{D.5}$$

Using similar steps as in Eaton and Kortum (2002), this relationship between the price and $mc_t(r)$ as well as the Fréchet distribution of $z_t^{\omega}(r)$ can be used to obtain the share of location s's spending on varieties from location r as

$$\pi_t(r,s) = \frac{T_t(r)\left[mc_t(r)\,\varsigma_t(r,s)\right]^{-\theta}}{\int_S T_t(u)\left[mc_t(u)\,\varsigma_t(u,s)\right]^{-\theta} du} \tag{D.6}$$

while, by the definition of the CES price index,

$$P_t(s) = \Gamma\left(1 - \frac{\rho}{(1-\rho)\theta}\right)^{-\frac{1-\rho}{\rho}}\left[\int_S T_t(r)\left[mc_t(r)\,\varsigma_t(r,s)\right]^{-\theta} dr\right]^{-\frac{1}{\theta}}. \tag{D.7}$$

Firms' Cobb–Douglas technology in labor and land and the local redistribution of land rents imply that land rents at location r are proportional to labor income:

$$R_t(r) = \frac{\xi - \mu\xi - \gamma_1}{\mu\xi + \gamma_1} w_t(r)\,\bar{L}_t(r) \tag{D.8}$$

and therefore market clearing conditions can be written in the form

$$w_t(r)\,H_t(r)\,\bar{L}_t(r) = \int_S \pi_t(r,s)\,w_t(s)\,H_t(s)\,\bar{L}_t(s)\,ds. \tag{D.9}$$

For a given initial spatial distribution of productivity $\tau_0(\cdot)$, equations (D.1) to (D.9) characterize the equilibrium spatial distribution of productivity, innovation, wages, land rents, trade flows and population levels in any period $t \in \{0, 1, ...\}$. Desmet et al. (2018) show that the equilibrium of the model exists and is unique if

$$\frac{\alpha}{\theta} + \frac{\gamma_1}{\xi} < \lambda + 1 - \mu + \Omega \tag{D.10}$$

which is a condition that, intuitively, requires that agglomeration forces are weaker than congestion forces. Following the same steps as in Desmet et al. (2018), it is straightforward to prove that condition (D.10) is also sufficient for the existence and uniqueness of the equilibrium in this model. Moreover, condition (D.10) holds under the values of structural parameters chosen in the calibration. Hence, one can be sure that the equilibrium of the DNR model found in the simulation is indeed the only equilibrium.

In the model of Desmet et al. (2018), the equilibrium growth rate of the world economy is increasing in world population. I assume that world population grows in the current model, which, everything else fixed, should then lead to accelerating growth. To avoid this, I follow footnote 11 in Desmet et al. (2018) and assume that the cost of innovation, ν_t, is linear in world population: $\nu_t = \nu \bar{L}_t$. In this case, Desmet et al. (2018) show that the growth rate is still a function of the population distribution, but not a function of $\bar{L}_t$.

D.2 Taking the DNR model to the data

This section describes how I take the DNR model to the data. Similarly to the model of Section 3, I treat each 20 by 20 arc minute grid cell of the U.S. as one location. One additional location represents Europe. I calculate the supply of land, $H_t(r)$, at each U.S. location in the same way as in the model of Section 3. I set the supply of land in Europe, $H_t(e)$, to $10,000$, as the land area of the European continent is approximately 10,000 times the land area of a grid cell. I also follow the model of Section 3 by setting the length of one period to five years.

Next, I need to choose the values of structural parameters, the distribution of shipping costs, amenities, initial productivity at time $t = 0$, and moving costs. I discuss each of these below.

Calibration of structural parameters. The scale parameter of firms' productivity draws, θ, equals the trade elasticity in the DNR model, which I set to eight, in line with how I calibrate the model of Section 3. I choose the share of land in production to equal 0.2, guided by Desmet and Rappaport (2017). This implies $\mu = 0.8$. Just like Desmet et al. (2018), I calibrate the level parameter of innovation costs, ν, to match the growth rate of the economy. In particular, I choose the value of ν for which the annual growth rate in U.S. real GDP per capita between 1800 and 1820 matches the growth rate in the data, 0.5% (Weiss, 1992). Unfortunately, the data do not provide any guidance on the calibration of the remaining structural parameters. As a result, I set them equal to their baseline calibrated values in Desmet et al. (2018). This implies setting $\alpha = 0.06$, $\xi = 125$, $\gamma_1 = 0.319$, $\gamma_2 = 0.99246$, $\Omega = 0.5$ and $\lambda = 0.32$.

Shipping costs. I calculate shipping costs across locations in the same way as in the model of Section 3. In particular, I borrow the relative costs of shipping via rail, water or wagon from Fogel (1964) and Donaldson and Hornbeck (2016). Just like in the calibration of the model of Section 3, I calibrate parameter ϕ_B, which drives the level of per-distance shipping costs between U.S. coastal locations and Europe, to match the U.S. export to GDP ratio in 1790. The level of per-distance shipping costs within the U.S. are driven by two separate parameters in the model of Section 3: parameter ϕ_F for the farm sector, and parameter ϕ_N for the non-farm sector. The DNR model features only one sector, hence only one parameter ϕ. In the baseline calibration of the DNR model, I set $\phi = \phi_F$, but the results are robust to setting $\phi = \phi_N$ instead.

Amenities and initial productivity. In Desmet et al. (2018), fundamental amenities relative to initial utility $\frac{\bar{a}(r)}{u_0(r)}$ and initial productivity levels $\tau_0(r)$ are calibrated such that the model matches the period-0 distribution of population and nominal income across locations. This procedure cannot be applied in my setting, as there are no consistent data on nominal income at the grid cell level for the pre-Civil War United States. As a result, I make the assumption that every location features the same level of amenities, which I normalize to $\bar{a}(r) = 1$. It is an extreme assumption, but it is consistent with the model of Section 3, in which I also abstract from amenity differences across locations. As the primary aim of the DNR model is that it serves as a benchmark to which my model's fit can be compared, it is in fact reassuring that the two models share this feature.

In Desmet et al. (2018), initial utility levels $u_0(r)$ are chosen to match country-level subjective wellbeing data, with the assumption that $u_0(r)$ do not vary within countries. I adopt this assumption, setting $u_0(r) = u_0$ for every U.S. location and normalizing u_0 to one. For Europe, I calibrate the value of initial utility $u_0(e)$ to match Europe's real GDP in 1790, which, by the assumption that $\bar{a}(e) = 1$, also equals $u_0(e)$.

Once $\frac{\bar{a}(r)}{u_0(r)}$ are known for every r, equations (D.1) to (D.9) can be used to back out the distribution of initial productivities $\tau_0(\cdot)$ that rationalize the initial distribution of population in 1790, $\bar{L}_0(\cdot)$. In particular, we have the following lemma.

Lemma 4. *Given $\frac{\bar{a}(\cdot)}{u_0(\cdot)}$, $H_0(\cdot)$, $\varsigma_0(\cdot,\cdot)$ and $\bar{L}_0(\cdot)$, the spatial distribution of initial productivity levels $\tau_0(\cdot)$ can be obtained as the solution to the equation*

$$\left[\frac{\bar{a}(r)}{u_0(r)}\right]^{-\frac{\theta(1+\theta)}{1+2\theta}} \tau_0(r)^{-\frac{\theta}{1+2\theta}} H_0(r)^{\frac{\theta}{1+2\theta}} \bar{L}_0(r)^{\lambda\theta - \frac{\theta}{1+2\theta}\left[\alpha - 1 + \left[\lambda + \frac{\gamma_1}{\xi} - (1-\mu)\right]\theta\right]} =$$

$$\kappa_1 \int_S \left[\frac{\bar{a}(s)}{u_0(s)} \right]^{\frac{\theta^2}{1+2\theta}} \tau_0(s)^{\frac{1+\theta}{1+2\theta}} H_0(s)^{\frac{\theta}{1+2\theta}} \varsigma_0(s)^{-\theta} \bar{L}_0(s)^{1-\lambda\theta+\frac{1+\theta}{1+2\theta}\left[\alpha-1+\left[\lambda+\frac{\gamma_1}{\xi}-(1-\mu)\right]\theta\right]} ds$$

where $\kappa_1 = \left[\frac{\mu\xi+\gamma_1}{\xi}\right]^{-\left(\mu+\frac{\gamma_1}{\xi}\right)\theta} \mu^{\mu\theta} \left[\frac{\xi\nu}{\gamma_1}\right]^{-\frac{\gamma_1\theta}{\xi}} \Gamma\left(1-\frac{\rho}{(1-\rho)\theta}\right)^{\theta\frac{1-\rho}{\rho}}$.

Proof. Combining equations (D.1), (D.2) and (D.4) to (D.9) yields the result. □

Lemma 4 provides me with an initial productivity level for every location that was part of the U.S. in 1790. I assign the productivity of the nearest location to locations that became part of the U.S. due to the political border changes after 1790.

Moving costs. I follow the procedure in Desmet et al. (2018) to back out location-specific moving costs $m_2(r)$. In particular, I choose $m_2(r)$ such that I exactly match the change in location r's population between 1790 and 1795. For U.S. grid cells, population data are available for the census years of 1790 and 1800 but not for 1795. Hence, I assume that the growth rate of population in any cell was the same between 1790 and 1795 as between 1795 and 1800. Desmet et al. (2018) show that, whenever condition (D.10) holds, this procedure identifies a unique set of moving costs that rationalize the changes in population. Just like with productivity, I assign the moving cost of the nearest locations to locations that became part of the U.S. after 1790.

D.3 Simulation

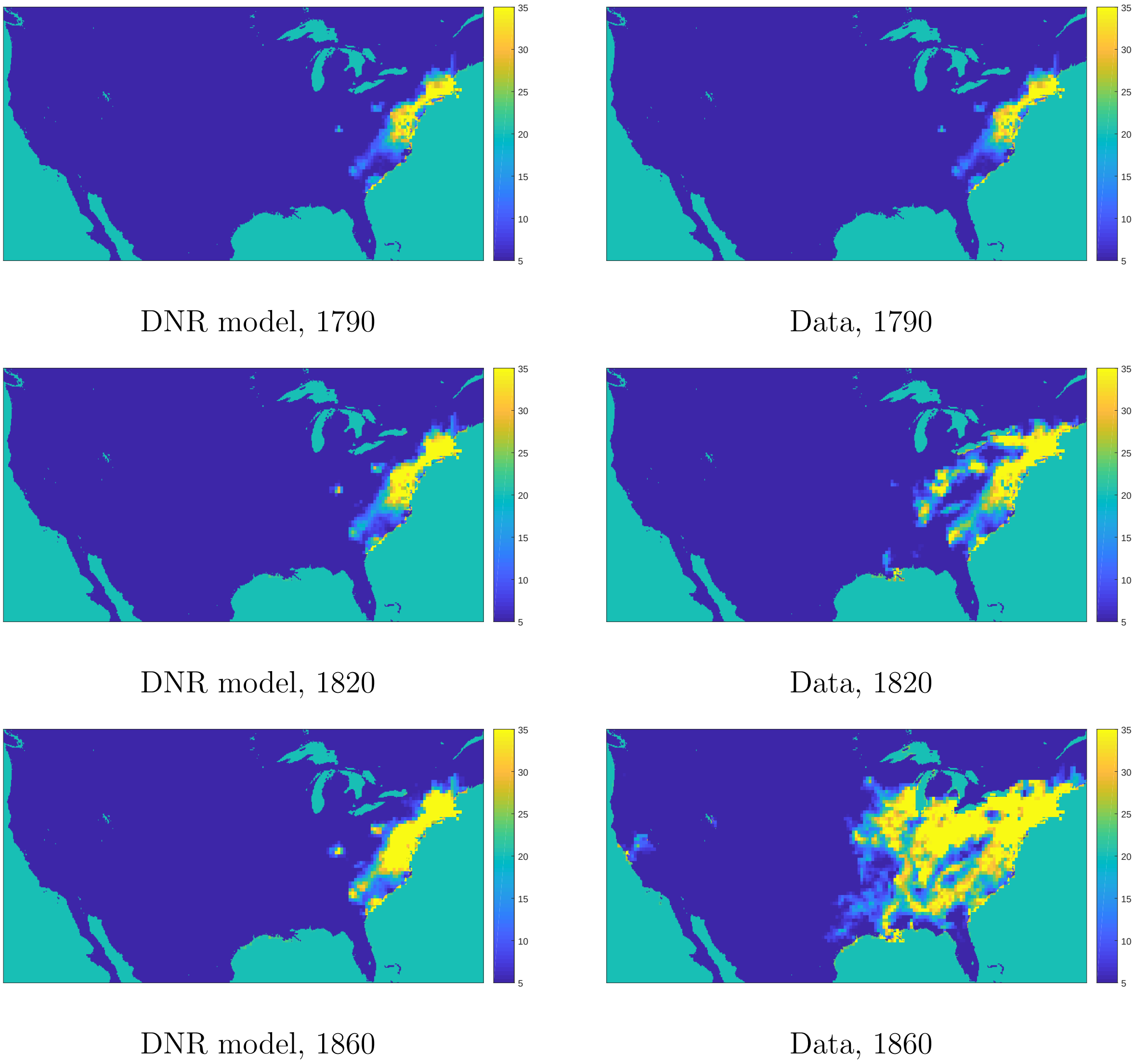
Armed with values of the model's fundamentals, I simulate the DNR model for 14 subsequent periods, corresponding to the years 1795, 1800, ..., 1860.[15] During the simulation, I change world population $\bar{L}_t$ (sum of the population of the U.S. and Europe), U.S. land $H_t(r)$ and shipping costs $\varsigma_t(r,s)$ as they change in the data. In particular, I feed the construction of railroads and canals into the model through changes in $\varsigma_t(r,s)$. To simulate the model, I use the iterative procedure proposed in Desmet et al. (2018).

D.4 Results

Figure 11 compares the evolution of the population distribution in the DNR model to the data. The figure shows that the DNR model is unable to replicate the first-order feature of the data: the substantial reallocation of U.S. population to the West. The intuition for this is as follows. In the DNR model, locations with the highest population density (and therefore the largest markets) offer the highest returns to innovation. As population density is initially the highest in the Northeast, these locations innovate the most, which reinforces their productivity advantage over the rest of the United States.

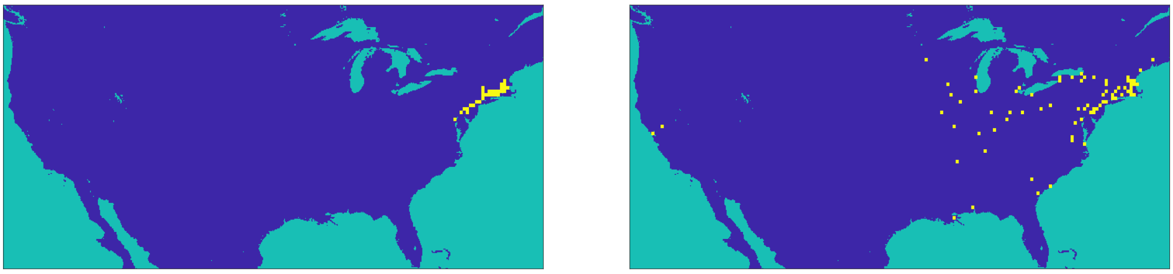
[15]The model does not need to be simulated for 1790, as it matches the 1790 population data exactly.

Figure 11: Population distribution: DNR model versus data



The three left panels of the figure present population (number of inhabitants) per square mile in each 20 by 20 arc minute grid cell of the U.S. in the DNR model in 1790, 1820 and 1860, respectively. The three right panels present the same in the data. The scale is bottom-coded at 5 and top-coded at 35 for visibility. Source: U.S. census.

Figure 12: City locations by 1860: DNR model versus data



DNR model

Data

In the left map, yellow 20 by 20 arc minute grid cells are those that have a city in them in 1860 in the DNR model. In the right map, yellow grid cells are those that have a city in them forming by 1860 in the data. Source: U.S. census.

Changing political borders and the expanding transport infrastructure in the West mitigate this effect, but are quantitatively unable to overturn it.

A key difference relative to the DNR model is that innovation is not a function of local population density in the model of Section 3. Instead, innovation depends on the size of the location's effective hinterland (Proposition 1). This feature of the model helps locations in the Midwest, with access to a large farm hinterland especially after the construction of railroads, to grow fast and attract a substantial fraction of U.S. population by 1860. This force, coming from the interaction between the farm and non-farm sectors, is absent from one-sector models such as the DNR model.

In line with this, the DNR model seems unable to rationalize the formation of cities in the Midwest. Although it is not obvious how to define the notion of a city in a one-sector model, I can select the cells with the highest population density in 1860, such that their total population equals total city population in the data. This gives me a total of 30 city cells in 1860. Figure 12 presents the location of these cells and contrasts them with the locations of 1860 cities in the data. As expected, the DNR model features a single cluster of cities in the Northeast, unlike the data and the model of Section 3.

In conclusion, the one-sector framework of Desmet et al. (2018) is unable to replicate the striking westward expansion of population and the process of city formation. The reason the model of Section 3 does substantially better seems to be that it gives a role to farm hinterlands in the growth of the non-farm sector. This interaction between the two sectors allows cities to form at Western locations with access to a large farm hinterland, which in turn can explain both the shift of U.S. economic activity to the West and the geographic patterns of city formation prior to the Civil War. In other words, the comparison of my model to the DNR framework further highlights the quantitative relevance of the hinterland hypothesis.

D.5 Robustness

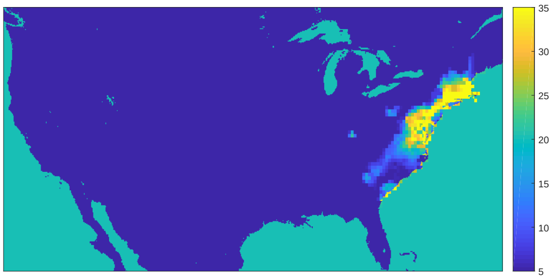
The simulation procedure of Appendix D.3 matches world population every period, but not necessarily U.S. population every period. In particular, U.S. population grows somewhat less in the model than in the data, which suggests that the frictions to moving from Europe to the U.S. might have become weaker over time. To see if my results are robust to this discrepancy between the model and the data, I simulate an alternative version of the DNR model, which I call "DNR (match total U.S. population)." In this version of the model, I let $m_2(e)$ change over time. As only the ratio of moving costs matters for the equilibrium, this procedure is equivalent to letting the cost of entering any U.S. location change to the same extent. I choose the value of $m_{2t}(e)$ in every period t such that the model matches U.S. population at t.

Figure 13 presents the evolution of the population distribution in the "DNR (match total U.S. population)" model. The figure shows that matching total U.S. population does not change the overall picture: the one-sector DNR model cannot replicate the westward movement of U.S. population. This is not surprising as the ratio of two locations' innovation rates depends on the ratio of their population density in the DNR model, not on absolute population levels (Desmet et al., 2018).

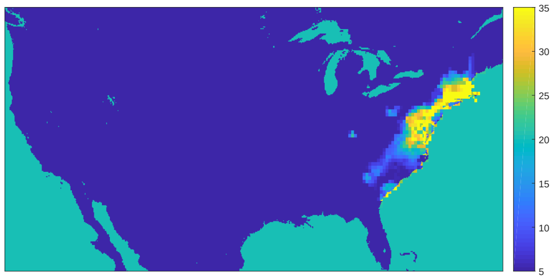
As a second robustness check, I also simulate an alternative version of the DNR model in which productivity does not grow over time. This alternative model, which I call "DNR (no productivity growth)," can be thought of as a one-sector static quantitative model of economic geography, simulated over a sequence of time periods over which total world population, U.S. land and shipping costs change as in the data.

Figure 14 shows that this model also fails to replicate the westward movement of population. This is not surprising as, unlike the model of Section 3, this alternative model also lacks the relationship between farm hinterland size and productivity growth. Figure 14 confirms that absent this relationship, cities do not start to develop and attract a substantial fraction of U.S. population to the West.

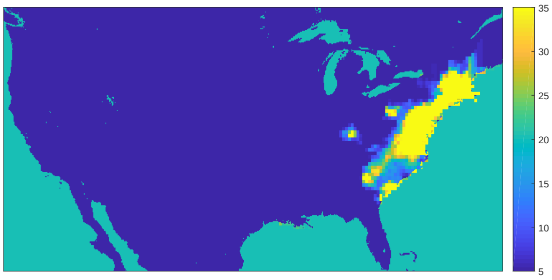
Figure 13: Population distribution: DNR (match total U.S. population) versus data



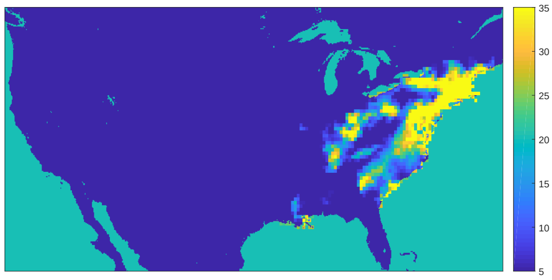
DNR model, 1790



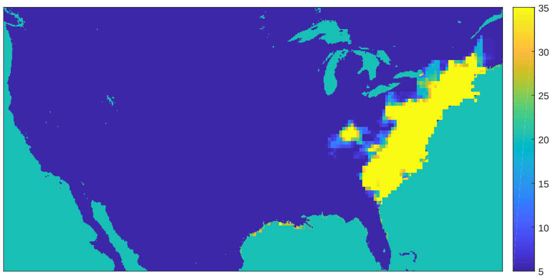
Data, 1790



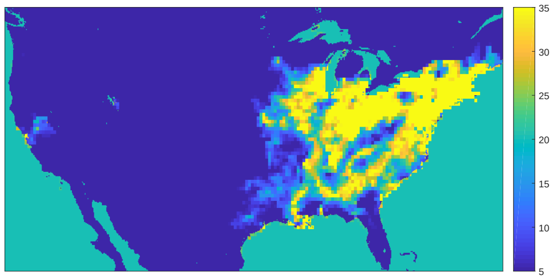
DNR model, 1820



Data, 1820



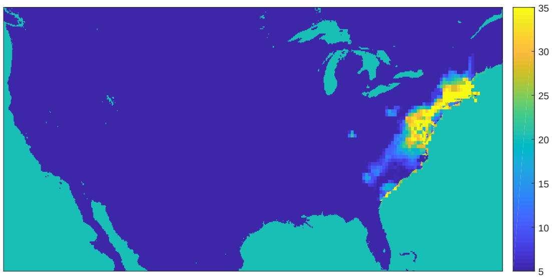
DNR model, 1860



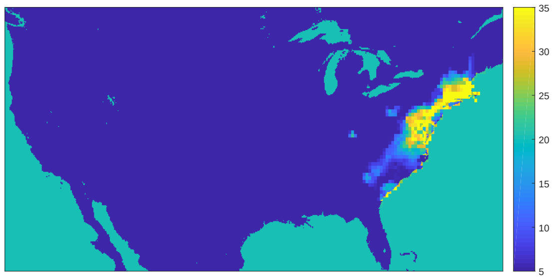
Data, 1860

The three left panels of the figure present population (number of inhabitants) per square mile in each 20 by 20 arc minute grid cell of the U.S. in the DNR (match total U.S. population) model in 1790, 1820 and 1860, respectively. The three right panels present the same in the data. The scale is bottom-coded at 5 and top-coded at 35 for visibility. Source: U.S. census.

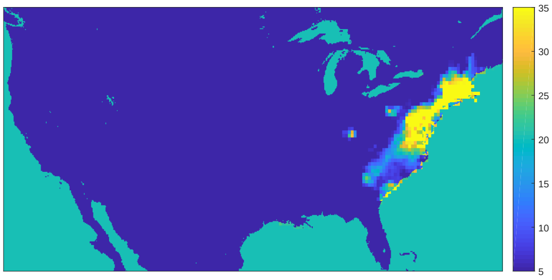
Figure 14: Population distribution: DNR (no productivity growth) versus data



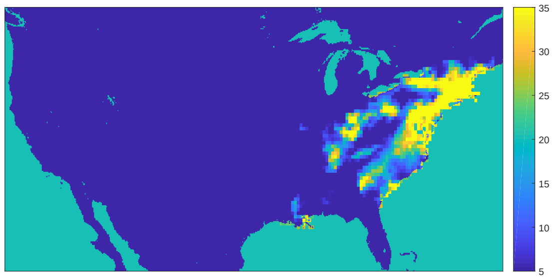
DNR model, 1790



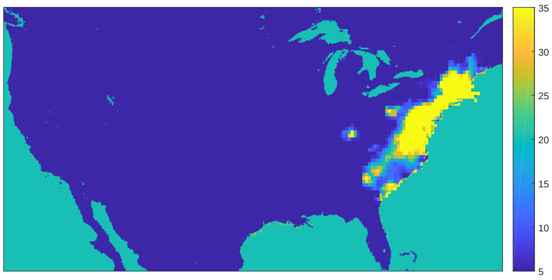
Data, 1790



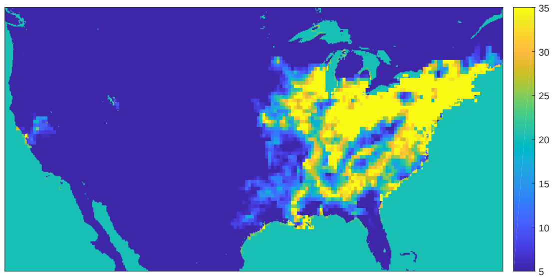
DNR model, 1820



Data, 1820



DNR model, 1860



Data, 1860

The three left panels of the figure present population (number of inhabitants) per square mile in each 20 by 20 arc minute grid cell of the U.S. in the DNR (no productivity growth) model in 1790, 1820 and 1860, respectively. The three right panels present the same in the data. The scale is bottom-coded at 5 and top-coded at 35 for visibility. Source: U.S. census.

E Additional tables and figures

Table 14: Discontinuity in the population growth rate of settlements

Dependent variable: ln(growth)				
	Threshold between cities and towns			
	(1)	(2)	(3)	(4)
	10,000	6,000	8,000	15,000
ln(size)	-0.02	0.05**	0.01	-0.01
	(0.02)	(0.02)	(0.02)	(0.03)
city	0.12**	-0.05	0.05	0.10
	(0.05)	(0.04)	(0.04)	(0.06)
No. of observations	371	371	371	371
R^2	0.03	0.01	0.01	0.02

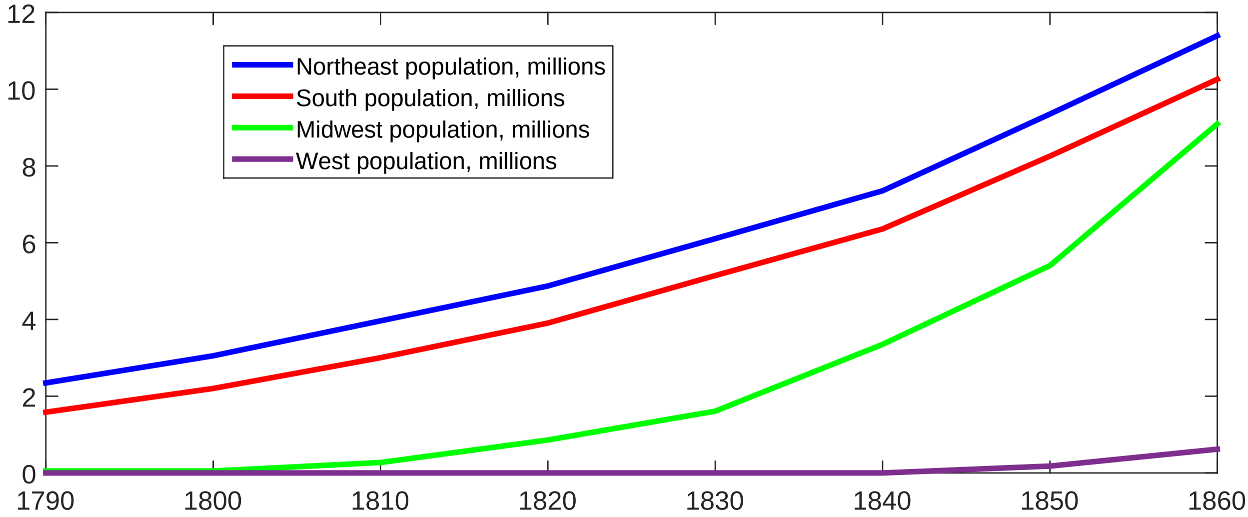
This table presents the results of regressing log settlement population growth on log settlement size and a variable indicating whether the settlement is a city. The unit of observation is a settlement (city or town) in the United States in a given census year (1790, 1800, 1810, 1820, 1830, 1840 or 1850). The dependent variable, *ln(growth)*, is the log of the settlement's population growth over the subsequent decade. Independent variable *ln(size)* is the log of the settlement's population. Independent variable *city* is a dummy variable equal to 1 if the settlement is classified as a city in the given year. A settlement is classified as a city if it exceeds a given population threshold. The columns of the table correspond to different population thresholds. For instance, column (1) classifies every settlement above 10,000 inhabitants in the given year as a city. Heteroskedasticity-robust standard errors in parentheses. *: significant at 5%; **: significant at 1%. Source: U.S. census.

Table 15: Large regions' shares in total U.S. population, 1860

Region	Model	Data
Northeast	41.7%	36.3%
South	26.4%	32.7%
Midwest	31.1%	29.0%
West	0.8%	2.0%

The first column of this table presents the share of 1860 U.S. population living in either the four large regions (Northeast, South, Midwest and West) in the baseline model simulation. The second column presents the same population shares in the data. Source: U.S. census.

Figure 15: Population of the four large U.S. regions

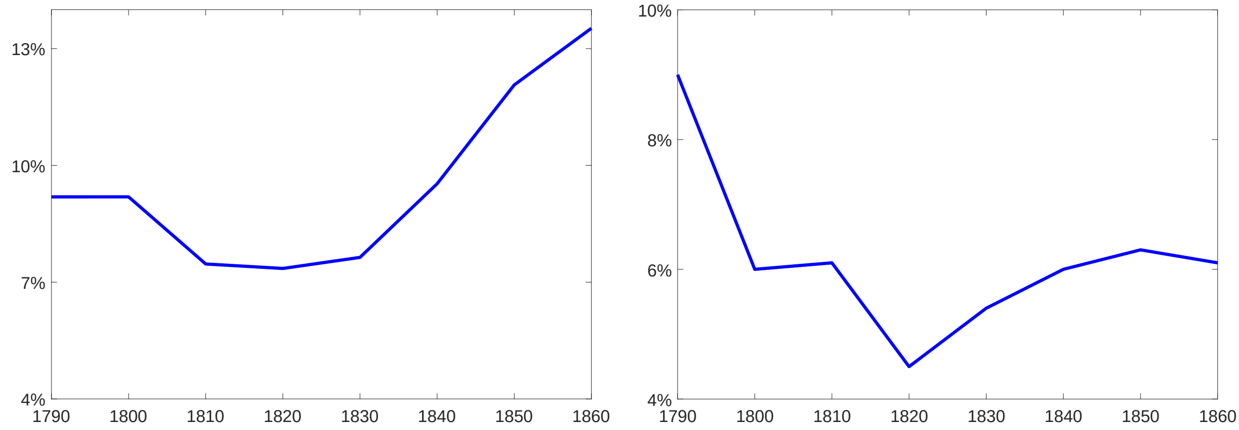


The four lines represent the population of the four large U.S. regions (Northeast, South, Midwest and West), measured in millions of inhabitants, in each census year. Source: U.S. census.

Figure 16: U.S. cities forming between 1790 and 1860

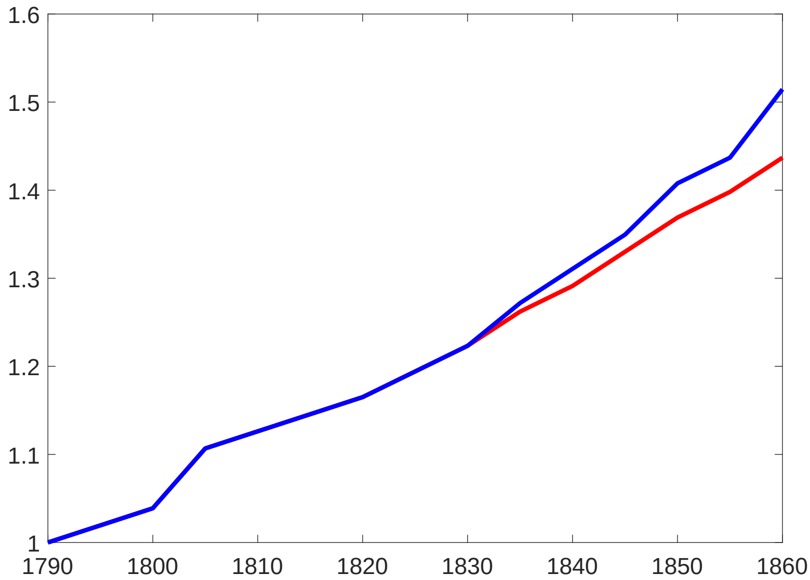
Each red dot in the map represents a city (a settlement above 10,000 inhabitants) that came into existence between 1790 and 1860 in the United States. The West has been cut for visibility. Source: U.S. census.

Figure 17: Exports to U.S. GDP: model versus data



The left graph shows the ratio of exports, measured in U.S. prices, to U.S. GDP in each census year in the baseline model simulation. The right graph shows the same object in the data. Source: Lipsey (1994).

Figure 18: U.S. real GDP per capita, baseline simulation (blue) versus no railroads (red)



The blue line corresponds to U.S. real GDP per capita relative to 1790 in the baseline model simulation. The red line corresponds to U.S. real GDP per capita relative to 1790 in the counterfactual with no railroads (Section 5.2).

References

- Allen, T. and Arkolakis, C. (2014): Trade and the topography of the spatial economy. *Quarterly Journal of Economics*, vol. 129(3), 1085–1140.
- Caselli, F. and Coleman, J. (2001): The U.S. structural transformation and regional convergence: A reinterpretation. *Journal of Political Economy*, vol. 109(3), 584–616.
- Desmet, K., Kopp, R., Kulp, S., Nagy, D., Oppenheimer, M., Rossi-Hansberg, E. and Strauss, B. (2021): Evaluating the economic cost of coastal flooding. *American Economic Journal: Macroeconomics*, vol. 13(2), 444–486.
- Desmet, K., Nagy, D., and Rossi-Hansberg, E. (2018): The geography of development. *Journal of Political Economy*, vol. 126(3), 903–983.
- Desmet, K. and Rappaport, J. (2017): The settlement of the United States, 1800-2000: The long transition to Gibrat's Law. *Journal of Urban Economics*, vol. 98, 50–68.
- Desmet, K. and Rossi-Hansberg, E. (2013): Urban accounting and welfare. *American Economic Review*, vol. 103, 2296–2327.
- Donaldson, D. and Hornbeck, R. (2016): Railroads and american economic growth: A "market access" approach. *Quarterly Journal of Economics*, vol. 131(2), 799–858.
- Eaton, J. and Kortum, S. (2002): Technology, geography, and trade. *Econometrica*, vol. 70(5), 1741–1779.
- Fogel, R. (1964): Railroads and American economic growth: Essays in econometric history. The Johns Hopkins Press.
- Krugman, P. (1991): Increasing returns and economic geography. *Journal of Political Economy*, vol. 99(3), 483–499.
- Lipsey, R. (1994): U.S. foreign trade and the balance of payments, 1800-1913. *NBER Working Paper* 4710.
- Nagy, D. (2022): Trade and urbanization: Evidence from Hungary. *American Economic Journal: Microeconomics*, vol. 14(3), 733–790.
- Starrett, D. (1978): Market allocations of location choice in a model with free mobility. *Journal of Economic Theory*, vol. 17, 21–37.
- Weiss, T. (1992): U.S. labor force estimates and economic growth, 1800-1860. In *American Economic Growth and Standards of Living before the Civil War*, 19–78. University of Chicago Press.