Online Appendix: Mechanics of Spatial Growth

The online appendix includes detailed theoretical derivations and proofs, additional quantitative results, and detailed data descriptions described in the paper.

A Proofs and Derivations of Idea Diffusion

In this appendix, we derive the idea diffusion process with a generic source distribution of insights. We then endogenize the source distribution as a result of idea diffusion from migrants and sellers.

A.1 Law of Motion of the Stock of Knowledge

In this section of the appendix, we derive the law of motion of knowledge with a generic source distribution of insights. .

We start by providing a proof of Proposition 1 and then derive the law of motion with idea flows after specifying the external source of ideas. In this section we use uppercase letters for random variables, and lowercase letters for their realized values.

Proposition 1 *Under Assumption 1, between t and $t+1$, the probability that the best new idea has productivity no greater than q—namely, $F_t^{best\ new}(q) = \Pr[\text{all new ideas are no greater than } q]$— is given by*

$$F_t^{best\ new}(q) = \exp\left(-\alpha_t q^{-\theta} \int_0^\infty x^{\rho\theta} dG_t(x)\right)$$

in the limiting case when $\bar{z} \to 0$.

Proof: *For any new idea that arrives between time t and $t+1$, the probability at time t that its productivity is no greater than q is given by*

$$\begin{aligned}
F_t^{new}(q) &= \Pr[ZQ'^{\rho} \leq q] \\
&= \int_0^\infty \Pr\left[Z \leq \frac{q}{Q'^{\rho}} \middle| q'\right] dG_t(q') \\
&= \int_0^{(q/\bar{z})^{1/\rho}} \Pr\left[Z \leq \frac{q}{Q'^{\rho}} \middle| q'\right] dG_t(q') + \int_{(q/\bar{z})^{1/\rho}}^\infty \Pr\left[Z \leq \frac{q}{Q'^{\rho}} \middle| q'\right] dG_t(q') \\
&= \int_0^{(q/\bar{z})^{1/\rho}} \Pr\left[Z \leq \frac{q}{Q'^{\rho}} \middle| q'\right] dG_t(q'). \\
&= \int_0^{(q/\bar{z})^{1/\beta}} H\left(\frac{q}{q'^{\beta}}\right) dG_t(q'),
\end{aligned}$$

where the fourth equality follows from the fact that $\Pr\left[Z \leq \frac{q}{Q'^{\beta}} \middle| Q' > (q/\bar{z})^{1/\beta}\right] = \Pr\left[Z \leq \bar{z}\right] = 0$. Using Assumption 1 a), on the functional form of $H(\cdot)$, we obtain

$$F_t^{new}(q) = \int_0^{(q/\bar{z})^{1/\rho}} \left[1 - \left(\frac{q/\bar{z}}{q'^{\rho}}\right)^{-\theta}\right] dG_t(q').$$

Note that in order to derive this expression, we do not need to specify the source distribution of the insights. Assumption 1 c) implies that between t and $t+1$, the probability that the best new idea has productivity no greater than q is given by

$$\begin{aligned}
&F_t^{best\ new}(q) \\
&= \Pr\left[\text{all new ideas are no greater than } q\right] \\
&= \sum_{s=0}^{\infty} \Pr[\text{\# new ideas = s}] \cdot \Pr\left[\text{all new ideas are no greater than } q \mid \text{\# new ideas = s}\right] \\
&= \sum_{s=0}^{\infty} \frac{(\alpha_t \bar{z}^{-\theta})^s e^{-(\alpha_t \bar{z}^{-\theta})}}{s!} \cdot F_t^{new}(q)^s \\
&= \underbrace{\sum_{s=0}^{\infty} \frac{\left[\alpha_t \bar{z}^{-\theta} F_t^{new}(q)\right]^s \cdot e^{-(\alpha_t \bar{z}^{-\theta}) F_t^{new}(q)}}{s!}}_{=1} \cdot e^{-(\alpha_t \bar{z}^{-\theta})(1-F_t^{new}(q))},
\end{aligned}$$

and therefore we obtain that

$$F_t^{best\ new}(q) = e^{-(\alpha_t \bar{z}^{-\theta})(1-F_t^{new}(q))}.$$

In order to characterize the probability distribution of the best new ideas, we hold α_t constant and investigate the limiting case where $\bar{z} \to 0$. We then have that

$$\begin{aligned}
\lim_{\bar{z}\to 0} \alpha_t \bar{z}^{-\theta}(1 - F_t^{new}(q)) &= \lim_{\bar{z}\to 0} \alpha_t \bar{z}^{-\theta} \left(1 - \int_0^{(q/\bar{z})^{1/\rho}} \left[1 - \left(\frac{q/\bar{z}}{q'^{\rho}}\right)^{-\theta}\right] dG_t(q')\right) \\
&= \lim_{\bar{z}\to 0} \alpha_t \bar{z}^{-\theta} \left(1 - G_t\left((\frac{q}{\bar{z}})^{\frac{1}{\rho}}\right) + \int_0^{(q/\bar{z})^{1/\rho}} \left[\left(\frac{q/\bar{z}}{q'^{\rho}}\right)^{-\theta}\right] dG_t(q')\right) \\
&= \alpha_t \lim_{\bar{z}\to 0} \bar{z}^{-\theta} \left[1 - G_t\left((\frac{q}{\bar{z}})^{\frac{1}{\rho}}\right)\right] + \alpha_t \lim_{\bar{z}\to 0} \bar{z}^{-\theta} \int_0^{(q/\bar{z})^{1/\rho}} \left[\left(\frac{q/\bar{z}}{q'^{\rho}}\right)^{-\theta}\right] dG_t(q') \\
&= \alpha_t \lim_{\bar{z}\to 0} \bar{z}^{-\theta} \left[1 - G_t\left((\frac{q}{\bar{z}})^{\frac{1}{\rho}}\right)\right] + \alpha_t \int_0^{\infty} \left(\frac{q}{q'^{\rho}}\right)^{-\theta} dG_t(q'),
\end{aligned}$$

where the first term on the right-hand side is zero by Assumption 1 d). In the limiting case when $\bar{z} \to 0$, the expression is equal to the second term only, which is $-\alpha_t q^{-\theta} \int_0^{\infty} x^{\rho\theta} dG_t(x)$.

Henceforth, we assume $\bar{z} \to 0$ and focus on the limiting case. The best new idea then follows

$$F_t^{best\ new}(q) = \exp\left(-\alpha_t q^{-\theta} \int_0^{\infty} x^{\rho\theta} dG_t(x)\right).$$

The productivity of the economy depends on the frontier of knowledge, $F_t(q)$. The frontier of knowledge denotes the fraction of varieties whose best producer has productivity no greater than q. In a probabilistic sense, $F_t(q)$ is also the probability that the best productivity for a specific variety is no greater than q at time t.

Proposition 2. *Assume that the initial frontier of knowledge at time* 0 *follows a Fréchet distribution given by* $F_0(q) = \exp(-A_0 q^{-\theta})$.

Imposing this assumption, then it follows that $F_t(\cdot)$ *is Fréchet at any* t *given by*

$$\begin{aligned} F_t(q) &= \exp\left[-\left(A_0 + \sum_{\tau=0}^{t-1} \alpha_\tau \int_0^\infty x^{\rho\theta} dG_\tau(x)\right) q^{-\theta}\right] \\ &= \exp\left(-A_t q^{-\theta}\right), \end{aligned}$$

where the law of motion for the knowledge stock is given by

$$A_{t+1} = A_t + \alpha_t \int_0^\infty x^{\rho\theta} dG_t(x).$$

Proof: *The frontier* $F_t(q)$ *changes from* t *to* $t+1$ *because some new ideas might have better productivity than the current best. At* $t+1$, *we then have*

$$\begin{aligned} F_{t+1}(q) &= \Pr\left[\textit{the best productivity is no greater than } q \textit{ at } t+1\right] \\ &= \Pr\left[\textit{the best productivity is no greater than } q \textit{ at } t\right] \cdot \\ &\quad \Pr[\textit{no new ideas greater than } q \textit{ between } t \textit{ and } t+1] \\ &= F_t(q) \cdot F_t^{\textit{best new}}(q) \\ &= F_0(q) \cdot \prod_{\tau=0}^{t} F_\tau^{\textit{best new}}(q), \end{aligned}$$

where the last line follows from iteration back to $t = 0$.

Assume that the initial distribution at time 0 *follows a Fréchet distribution; namely,*

$$F_0(q) = \exp(-A_0 q^{-\theta}).$$

Then it follows that $F_t(\cdot)$ *is Fréchet at any* t*:*

$$\begin{aligned} F_t(q) &= \exp\left[-\left(A_0 + \sum_{\tau=0}^{t-1} \alpha_\tau \int_0^\infty x^{\rho\theta} dG_\tau(x)\right) q^{-\theta}\right] \\ &= \exp\left(-A_t q^{-\theta}\right). \end{aligned}$$

It also follows that the law of motion of the knowledge stock is

$$A_{t+1} = A_t + \alpha_t \int_0^\infty x^{\rho\theta} dG_t(x).$$

As we can see from this equation, both the arrival rate of new ideas α_t and the learning pool $G_t(\cdot)$ matter for the evolution of A_t.

A.2 Migration and the Source Distribution of Insights

Assume that at time t in location n, when a new idea arrives, the insight from a randomly drawn person currently living in n is the insight component of the new idea. Then

$$\begin{aligned}
G_{n,t}(q') &= \Pr[\text{the insight component is no greater than } q'] \\
&= \sum_{i=1}^{N} \Pr[\text{the person with the insight lives in } i \text{ at } t] \cdot \\
&\quad\; \Pr[\text{the insight is no greater than } q' | \text{the person with the insight lives in } i \text{ at } t] \\
&= \sum_{i=1}^{N} s_{in,t} F_{i,t}(q'),
\end{aligned}$$

where $s_{in,t}$ is the share of households from location i living in location n. In particular, we denote by $\mu_{in,t}$ the fraction of households that relocate from from i to n. We then have

$$s_{in,t} = \frac{\mu_{in,t} L_{i,t}}{\sum_{h=1}^{N} \mu_{hn,t} L_{h,t}}$$

and

$$\int_0^\infty x^{\rho_\ell \theta} dG_t(x) = \Gamma(1-\rho_\ell) \sum_{i=1}^{N} s_{in,t} (A_{i,t})^{\rho_{\ell l}}.$$

Finally, the law of motion of the stock of knowledge with ideas from people is given by

$$A_{n,t+1} - A_{n,t} = \alpha_{n,t} \Gamma(1-\rho_\ell) \sum_{i=1}^{N} s_{in,t} (A_{i,t})^{\rho_\ell}.$$

A.3 Derivation of the Law of Motion of Knowledge with Ideas from Migrants and Sellers

Now we derive the law of motion of the knowledge stock with idea flows from both trade and migration.

We impose the following version of Assumption 1 to incorporate both sources of idea flows:

Assumption 1′

a) The same as Assumption 1 a)

b) The strength of idea diffusion, $\rho_m + \rho_l \in [0, 1)$, is strictly less than 1.

c) The same as Assumption 1 c)

d) The source distribution has a sufficiently thin tail such that for any monotonically decreasing sequence of $\bar{z}_n \to 0$, $\alpha_t \lim_{n\to\infty} \bar{z}_n^{-\theta} \left[1 - \int_{B(\bar{z}_n)} dG_t(q_\ell, q_m)\right] = 0$, where $B(\bar{z}) := \left\{(x_1, x_2) : \bar{z} x_1^{\rho_l} x_2^{\rho_m} < q\right\} \subset \mathbb{R}^2$. In addition, the integral $\int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} dG_t(q_\ell, q_m)$ exists.

Proposition 1′ *Under Assumption 1′, between t and $t+1$, the probability that the best new idea has productivity no greater than q—namely, $F_t^{best\ new}(q) = \Pr[$all new ideas are no greater than $q]$— is given by*

$$F_t^{best\ new}(q) = \exp\left(-\alpha_t q^{-\theta} \int_0^\infty \int_0^\infty \left(q_\ell^{\rho_\ell} q_m^{\rho_m}\right)^\theta dG_t(q_\ell, q_m)\right)$$

in the limiting case where $\bar{z} \to 0$.

Proof: *For any new idea that arrives between time t and $t+1$, the probability at t that its productivity is no greater than q is given by*

$$\begin{aligned}
F_t^{new}(q) &= \Pr[Z Q_\ell^{\rho_\ell} Q_m^{\rho_m} \leq q] \\
&= \int_{\mathbb{R}_+^2} \Pr\left[Z \leq \frac{q}{Q_\ell^{\rho_\ell} Q_m^{\rho_m}} \middle| q_\ell, q_m\right] dG_t(q_\ell, q_m) \\
&= \int_{B(\bar{z})} \Pr\left[Z \leq \frac{q}{Q_\ell^{\rho_\ell} Q_m^{\rho_m}} \middle| q_\ell, q_m\right] dG_t(q_\ell, q_m) + \int_{\mathbb{R}_+^2 \setminus B(\bar{z})} \Pr\left[Z \leq \frac{q}{Q_\ell^{\rho_\ell} Q_m^{\rho_m}} \middle| q_\ell, q_m\right] dG_t(q_\ell, q_m) \\
&= \int_{B(\bar{z})} \Pr\left[Z \leq \frac{q}{Q_\ell^{\rho_\ell} Q_m^{\rho_m}} \middle| q_\ell, q_m\right] dG_t(q_\ell, q_m),
\end{aligned}$$

where $B(\bar{z})$ is defined in Assumption 1′ d). Using Assumption 1′ a), we obtain

$$F_t^{new}(q) = \int_{B(\bar{z})} \left[1 - \left(\frac{q/\bar{z}}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta}\right] dG_t(q_\ell, q_m).$$

The probability that the best new idea has productivity no greater than q is the same as before: $F_t^{best\ new}(q) = e^{-(\alpha_t \bar{z}^{-\theta})(1 - F_t^{new}(q))}$. Consider a monotonically decreasing sequence of $\bar{z}_n \to 0$. We prove by the dominated convergence theorem that $\lim_{n\to\infty} \alpha_t \bar{z}_n^{-\theta}(1 - F_t^{new}(q)) = \alpha_t \int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} dG_t(q_\ell, q_m)$. The integral exists under Assumption 1′ d).

Define $g_n : \mathbb{R}_+ \to \mathbb{R}$,

$$\begin{aligned} g_n(q) &= \bar{z}_n^{-\theta}(1 - F_t^{new}(q)) \\ &= \bar{z}_n^{-\theta}\left(1 - \int_{B(\bar{z}_n)} \left[1 - \left(\frac{q/\bar{z}}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta}\right] dG_t(q_\ell, q_m)\right) \\ &= \bar{z}_n^{-\theta}\left[1 - \int_{B(\bar{z}_n)} dG_t(q_\ell, q_m)\right] + \int_{B(\bar{z}_n)} \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} dG_t(q_\ell, q_m) \\ &= \bar{z}_n^{-\theta}\left[1 - \int_{B(\bar{z}_n)} dG_t(q_\ell, q_m)\right] + \int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} \mathbf{1}_{B(\bar{z}_n)} dG_t(q_\ell, q_m). \end{aligned}$$

By Assumption 1' d), we have $\lim_{n\to\infty} \bar{z}_n^{-\theta}\left[1 - \int_{B(\bar{z}_n)} dG_t(q_\ell, q_m)\right] = 0$. *Since* $\forall q \geq 0$, $\forall n$,

$$\left|\left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} \mathbf{1}_{B(\bar{z}_n)}\right| \leq \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta},$$

and

$$\lim_{n\to\infty} \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} \mathbf{1}_{B(\bar{z}_n)} = \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta},$$

by the dominated convergence theorem, we have

$$\lim_{n\to\infty} \int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} \mathbf{1}_{B(\bar{z}_n)} dG_t(q_\ell, q_m) = \int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} dG_t(q_\ell, q_m),$$

so

$$\begin{aligned} \lim_{n\to\infty} g_n(q) &= \lim_{n\to\infty} \alpha_t \bar{z}_n^{-\theta}(1 - F_t^{new}(q)) \\ &= \lim_{n\to\infty} \bar{z}_n^{-\theta}\left[1 - \int_{B(\bar{z}_n)} dG_t(q_\ell, q_m)\right] + \lim_{n\to\infty} \int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} \mathbf{1}_{B(\bar{z}_n)} dG_t(q_\ell, q_m) \\ &= \int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} dG_t(q_\ell, q_m). \end{aligned}$$

Henceforth, we assume $\bar{z} \to 0$ *and focus on the limiting case. The best new idea then follows*

$$F_t^{best\ new}(q) = \exp\left(-\alpha_t q^{-\theta} \int \left(q_\ell^{\rho_\ell} q_m^{\rho_m}\right)^{\theta} dG_t(q_\ell, q_m)\right)$$

or, using the Riemann integral,

$$F_t^{best\ new}(q) = \exp\left(-\alpha_t q^{-\theta} \int_0^\infty \int_0^\infty \left(q_\ell^{\rho_\ell} q_m^{\rho_m}\right)^\theta dG_t(q_\ell, q_m)\right).$$

As in the previous section, in this section it follows that the frontier distribution $F_{n,t}(\cdot)$ follows a Fréchet distribution with location parameter $A_{n,t}$ and shape parameter θ, and the law of motion of $A_{n,t}$ is

$$A_{n,t+1} = A_{n,t} + \alpha_t \int_0^\infty \int_0^\infty \left(q_\ell^{\rho_\ell} q_m^{\rho_m}\right)^\theta dG_{n,t}(q_\ell, q_m).$$

Since we assume that q_l, q_m come from independent distributions, the law of motion becomes

$$A_{n,t+1} = A_{n,t} + \alpha_t \int_0^\infty q_\ell^{\theta\rho_\ell} dG_{n,t}^l(q_\ell) \int_0^\infty q_m^{\theta\rho_m} dG_{n,t}^m(q_m).$$

The first integral,

$$\int_0^\infty q_\ell^{\theta\rho_\ell} dG_{n,t}^l(q_\ell) = \Gamma(1-\rho_\ell) \sum_{i=1}^N s_{in,t} (A_{i,t})^{\rho_\ell},$$

is the same as in the previous section. The derivation of this term follows the previous section of this appendix. For the second integral, we assume learning from sellers as in Buera and Oberfield (2020). Namely, the insights from goods are randomly drawn from the set of goods sold locally. To simplify the notation, we omit intermediate goods in the derivation that follows. In this case,

$$\begin{aligned}
G_{n,t}^m(x) &= \sum_i \mathbb{P}\left[q_i \leq x,\ i \text{ is the lowest-cost supplier to } n \text{ at } t\right] \\
&= \sum_i \mathbb{P}\left[q_i \leq x, q_j \leq \frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} q_i\ \forall j\right] \\
&= \sum_i \int_0^x f_{i,t}(q) \left(\prod_{j\neq i} F_{j,t}\left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} q\right)\right) dq,
\end{aligned}$$

where $F_{i,t}(\cdot)$ and $f_{i,t}(\cdot)$ are the cumulative distribution function (CDF) and probability density function (PDF) of a Fréchet distribution with location parameter $A_{i,t}$ and shape parameter θ, respectively:

$$\begin{aligned}
F_{i,t}(q) &= \exp\left(-A_{i,t} q^{-\theta}\right), \\
f_{i,t}(q) &= A_{i,t}\theta q^{-\theta-1} \exp\left(-A_{i,t} q^{-\theta}\right).
\end{aligned}$$

Therefore,

$$
\begin{aligned}
G_{n,t}^{m}(x) &= \sum_{i} \int_{0}^{x} f_{i,t}(q) \left(\prod_{j \neq i} F_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} q \right) \right) dq \\
&= \sum_{i} \int_{0}^{x} A_{i,t}\theta q^{-\theta-1} \exp\left(-A_{i,t}q^{-\theta}\right) \exp\left(-\sum_{j\neq i} A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} q^{-\theta} \right) dq \\
&= \sum_{i} \int_{0}^{x} A_{i,t}\theta q^{-\theta-1} \exp\left(-\sum_{j} A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} q^{-\theta} \right) dq \\
&= \sum_{i} \frac{A_{i,t}\left(w_{i,t}\kappa_{ni}\right)^{-\theta}}{\sum_j A_{j,t}\left(w_{j,t}\kappa_{nj}\right)^{-\theta}} \exp\left(-\sum_{j} A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} x^{-\theta} \right) \\
&= \sum_{i} \pi_{ni,t} \exp\left(-\sum_{j} A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} x^{-\theta} \right).
\end{aligned}
$$

It follows that the second integral, which represents the learning from goods, is given by

$$
\begin{aligned}
\int_{0}^{\infty} q_m^{\theta\rho_m} dG_{n,t}^{m}(q_m) &= \int_{0}^{\infty} q_m^{\theta\rho_m} d \sum_{i} \pi_{ni,t} \exp\left(-\sum_{j} A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} q_m^{-\theta} \right) \\
&= \sum_{i} \pi_{ni,t} \int_{0}^{\infty} q_m^{\theta\rho_m} d \exp\left(-\sum_{j} A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} q_m^{-\theta} \right)
\end{aligned}
$$

Using change of variables, define $x = \sum_j A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} q_m^{-\theta}$*, and we have*

$$
\begin{aligned}
\int_{0}^{\infty} q_m^{\theta\rho_m} dG_{n,t}^{m}(q_m) &= \sum_{i} \pi_{ni,t} \int_{0}^{\infty} \sum_{j} A_{j,t}^{\rho_m} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta\rho_m} x^{-\rho_m} d \exp\left(-x\right) \\
&= \Gamma(1-\rho_m) \sum_{i} \pi_{ni,t} \left(\frac{A_{i,t}}{\pi_{ni,t}} \right)^{\rho_m}.
\end{aligned}
$$

Therefore, the law of motion of $A_{n,t}$ *is given by*

$$
A_{n,t+1} - A_{n,t} = \alpha_t \Gamma(1-\rho_\ell)\Gamma(1-\rho_m) \left[\sum_{i=1}^{N} s_{in,t}(A_{i,t})^{\rho_\ell} \right] \left[\sum_{i=1}^{N} \pi_{ni,t} \left(\frac{A_{i,t}}{\pi_{ni,t}} \right)^{\rho_m} \right].
$$

B Derivations of Demand and Supply Equilibrium Conditions

In this appendix, we provide detailed derivations of the trade shares, migration shares, and the solution to landowner consumption and investment decisions.

B.1 Derivation of the Trade Shares

Let Ω be the variety space and intermediate variety $\omega \in \Omega$. Let $p_{in,t}(\omega)$ be the price that firms in location i pay for good ω purchased from location n at time t. Then perfect competition implies

$$p_{in,t}(\omega) = \frac{\kappa_{in,t} x_{n,t}}{q(\omega)},$$

where $x_{n.t}$ is the unit cost of inputs to produce in location n. Since $\{q(\omega)\}_{\omega \in \Omega}$ are i.i.d., for all $\omega \in \Omega$, they have the same distribution. Let $H_{in,t}(p)$ be the cumulative distribution of prices, i.e., $H_{in,t}(p) = \mathbb{P}\left[p_{in,t}(\omega) \leq p\right]$. Then

$$\begin{aligned}
H_{in,t}(p) &= \mathbb{P}[p_{in,t}(\omega) \leq p] \\
&= \mathbb{P}\left[\frac{\kappa_{in,t} x_{n,t}}{q(\omega)} \leq p\right] \\
&= \mathbb{P}\left[q(\omega) \geq \frac{\kappa_{in,t} x_{n,t}}{p}\right] \\
&= 1 - \mathbb{P}\left[q(\omega) \leq \frac{\kappa_{in} x_{n,t}}{p}\right] \\
&= 1 - F_{n,t}\left(\frac{\kappa_{in,t} x_{n,t}}{p}\right) \\
&= 1 - \exp\left\{-A_{n,t}\left(\frac{\kappa_{in,t} x_{n,t}}{p}\right)^{-\theta}\right\},
\end{aligned} \tag{B.1}$$

where $F_{n,t}(\cdot)$ denotes the Fréchet distribution with scale parameter $A_{n,t}$ and shape parameter θ.

Let $\lambda_{in,t}$ be the fraction of goods purchased by location i from n. For location i to buy good ω from n, n must be the lowest-cost supplier among all locations. By the law of large numbers, we have

$$\begin{aligned}
\lambda_{in,t} &= \mathbb{P}\left[p_{in,t}(\omega) \leq \min_{h \in S \setminus \{i\}} p_{ih,t}(\omega)\right] \\
&= \int_0^\infty \mathbb{P}\left[\min_{h \in S \setminus \{i\}} p_{ih,t}(\omega) \geq p\right] dH_{in,t}(p) \\
&= \int_0^\infty \mathbb{P}\left[\bigcap_{h \in S \setminus \{i\}} \{p_{ih,t}(\omega) \geq p\}\right] dH_{in,t}(p) \\
&= \int_0^\infty \prod_{h \in S \setminus \{i\}} \mathbb{P}\left[p_{ih,t}(\omega) \geq p\right] dH_{in,t}(p) \\
&= \int_0^\infty \prod_{h \in S \setminus \{i\}} \left[1 - H_{ih,t}(p)\right] dH_{in,t}(p),
\end{aligned}$$

where the law of iterated expectation is used for the second equality and independence is used for the fourth equality.

Using the expression of price distribution derived in (B.1), we have

$$\lambda_{in,t} = \int_0^\infty \prod_{h\in S\setminus\{i\}} \exp\left\{-A_{h,t}\left(\frac{\kappa_{ih,t}x_{h,t}}{p}\right)^{-\theta}\right\} \exp\left\{-A_{n,t}\left(\frac{\kappa_{in,t}x_{n,t}}{p}\right)^{-\theta}\right\} A_{n,t}(\kappa_{in}x_{in,t})^{-\theta}dp^\theta$$
$$= A_{n,t}(\kappa_{in,t}x_{n,t})^{-\theta}\int_0^\infty \exp\left\{-\sum_{h=1}^N A_{h,t}(\kappa_{ih,t}x_{h,t})^{-\theta}p^\theta\right\}dp^\theta$$
$$= \frac{A_{n,t}(\kappa_{in,t}x_{n,t})^{-\theta}}{\sum_{h=1}^N A_{h,t}(\kappa_{ih,t}x_{h,t})^{-\theta}}.$$

B.2 Derivation of Gross Flows Equation

Let $\mu_{in,t}$ be the fraction of individuals who relocate from location i to location n at time t. By definition, we have

$$\mu_{in,t} = \mathbb{P}\left[\frac{\beta V_{n,t+1} - m_{in,t}}{\nu} + \epsilon_{n,t} \geq \max_{l\neq n}\left\{\frac{\beta V_{l,t+1} - m_{il,t}}{\nu} + \epsilon_{l,t}\right\}\right]$$
$$= \int_\infty^\infty \mathbb{P}\left[\frac{\beta V_{n,t+1} - m_{in,t}}{\nu} + x \geq \max_{l\neq n}\left\{\frac{\beta V_{l,t+1} - m_{il,t}}{\nu} + \epsilon_{l,t}\right\}\right] dM(x)$$
$$= \int_{-\infty}^\infty \mathbb{P}\left[\bigcap_{l\neq n}\left\{\frac{\beta V_{n,t+1} - m_{in,t}}{\nu} + x \geq \frac{\beta V_{l,t+1} - m_{il,t}}{\nu} + \epsilon_{l,t}\right\}\right] dM(x)$$
$$= \int_{-\infty}^\infty \prod_{l\neq n}\mathbb{P}\left[\frac{\beta V_{n,t+1} - m_{in,t}}{\nu} + x \geq \frac{\beta V_{l,t+1} - m_{il,t}}{\nu} + \epsilon_{l,t}\right] dM(x)$$
$$= \int_{-\infty}^\infty \prod_{l\neq n}\mathbb{P}\left[\epsilon_{l,t} \leq \frac{\beta(V_{n,t+1} - V_{l,t+1}) - (m_{in,t} - m_{il,t})}{\nu} + x\right] dM(x)$$
$$= \int_{-\infty}^\infty \prod_{l\neq n} M\left(\frac{\beta(V_{n,t+1} - V_{l,t+1}) - (m_{in,t} - m_{il,t})}{\nu} + x\right) dM(x),$$

where $M(\cdot)$ denotes the cumulative distribution function of a Gumbel Type I distribution.

Define $\bar{\epsilon}_{ln,t} = \frac{\beta(V_{n,t+1} - V_{l,t+1}) - (m_{in,t} - m_{il,t})}{\nu}$ with $\bar{\epsilon}_{nn,t} = 0$. Using this notation and the expression of $M(\cdot)$, we have

$$\mu_{in,t} = \int_{-\infty}^\infty \exp\{-e^{-x-\gamma}\}e^{-x-\gamma}\exp\{-e^{-x-\gamma}\sum_{l\neq n} e^{-\bar{\epsilon}_{ln,t}}\}dx$$
$$= \int_{-\infty}^\infty e^{-x-\gamma}\exp\{-e^{-x-\gamma}\sum_{l=1}^N e^{-\bar{\epsilon}_{ln,t}}\}dx.$$

Define $\Xi_{in} = \log(\sum_{l=1}^N e^{-\bar{\epsilon}_{ln,t}})$ and $y = x + \gamma - \Xi_{in}$. Then

$$\mu_{in,t} = \int_{-\infty}^\infty e^{-y-\Xi_{in}}e^{-e^{-y}}dy$$
$$= e^{\Xi_{in}}.$$

Finally, plugging in the expression of Ξ_{in}, we have

$$\mu_{in,t} = \frac{1}{\sum_{l=1}^{N} \exp\left\{\frac{\beta(V_{l,t+1}-V_{n,t+1})-m_{il,t}+m_{in,t}}{\nu}\right\}} = \frac{\exp(\beta V_{n,t+1} - m_{in,t})^{\frac{1}{\nu}}}{\sum_{l=1}^{N} \exp(\beta V_{l,t+1} - m_{il,t})^{\frac{1}{\nu}}}.$$

B.3 Landlord's Problem

The landlord's problem is defined as

$$\max_{\{C_{i,t},K_{i,t+1}\}_{t=0}^{\infty}} U = \sum_{t=0}^{\infty} \beta^t \log(C_{i,t}),$$
$$\text{s.t. } r_{i,t}K_{i,t} = P_{i,t}\left[C_{i,t} + K_{i,t+1} - (1-\delta)K_{i,t}\right] \text{ all } t,$$

where δ is the depreciation rate and $K_{i,0}$ is taken as given. Set up the Lagrangian equation,

$$\mathcal{L} = \left[\beta^t\{\log(C_{i,t}) + \lambda_t[r_{i,t}K_{i,t} - P_{i,t}(C_{i,t} + K_{i,t+1} - (1-\delta)K_{i,t})]\}\right], \tag{B.2}$$

where λ_t is the Lagrangian multiplier for the constraint in period t.

The first-order conditions for the problem are

$$\frac{1}{C_{i,t}} = \lambda_t P_{i,t}$$
$$\lambda_t P_{i,t} = \beta\left[\lambda_{t+1}[r_{i,t+1} + P_{i,t+1}(1-\delta)]\right].$$

Define $R_{i,t} = 1 - \delta + \frac{r_{i,t}}{P_{i,t}}$. Then eliminating λ_t yields the Euler equation,

$$\frac{1}{C_{i,t}} = \beta\left[R_{i,t+1}\frac{1}{C_{i,t+1}}\right], \tag{B.3}$$

together with the budget constraint

$$R_{i,t}K_{i,t} = C_{i,t} + K_{i,t+1}. \tag{B.4}$$

To solve this problem, we use the guess-and-verify strategy. We guess that $C_{i,t} = \varsigma R_{i,t}K_{i,t}$, where ς is a constant to be determined. Plugging in (B.4), we have

$$K_{i,t+1} = (1-\varsigma)R_{i,t}K_{i,t}. \tag{B.5}$$

Combining equations (B.5) and (B.3), we have

$$\frac{1}{\varsigma R_{i,t}K_{i,t}} = \beta\left[R_{i,t+1}\frac{1}{\varsigma R_{i,t+1}(1-\varsigma)R_{i,t}K_{i,t}}\right].$$

The undetermined coefficient method implies that $\varsigma = 1-\beta$. Hence, the consumption and saving policy functions are as follows:

$$\begin{aligned} C_{i,t} &= (1-\beta)[r_{i,t}/P_{i,t} + (1-\delta)]K_{i,t}, \\ K_{i,t} &= \beta[r_{i,t}/P_{i,t} + (1-\delta)]K_{i,t}. \end{aligned}$$

C Detrended Equilibrium Conditions and Balanced Growth Path

In this appendix we characterize the long-run growth rates of the equilibrium variables of the model at the balanced growth path. In what follows, we denote the long-run growth rate of any variable y_t by $(1+g_y)$, and we also refer to a variable with a "~" as a detrended. In particular, $\tilde{y}_t = y_t/\left(1+g_y\right)^t$.

We start with the evolution of the stock of knowledge. At the balanced growth path, $A_{n,t}$ for all n grow at a rate $1+g_A$. From the law of motion of the stock of knowledge, we have

$$A_{n,t+1} - A_{n,t} = \alpha_t\Gamma_\rho \sum_{i=1}^{N} s_{in,t}\left(A_{i,t}\right)^{\rho_l} \sum_{i=1}^{N} \lambda_{ni,t}\left(\frac{A_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m},$$

using Assumption 2 and after detrending the variables, we obtain

$$\tilde{A}_{n,t+1}\left(1+g_A\right)^{t+1} - \tilde{A}_{n,t}\left(1+g_A\right)^{t} = \alpha_0\left(1+g_\alpha\right)^t \Gamma_\rho \sum_{i=1}^{N} s_{in,t}\left(\tilde{A}_{i,t}\left(1+g_A\right)^t\right)^{\rho_l} \sum_{i=1}^{N} \lambda_{ni,t}\left(\frac{\tilde{A}_{i,t}\left(1+g_A\right)^t}{\lambda_{ni,t}}\right)^{\rho_m}$$

or

$$\tilde{A}_{n,t+1}\left(1+g_A\right) - \tilde{A}_{n,t} = \left(1+g_\alpha\right)^t\left(1+g_A\right)^{t(\rho_l+\rho_m-1)}\alpha_0\Gamma_\rho \sum_{i=1}^{N} s_{in,t}\left(\tilde{A}_{i,t}\right)^{\rho_l} \sum_{i=1}^{N} \lambda_{ni,t}\left(\frac{\tilde{A}_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m},$$

which then implies that the long-run growth rate of the stock of knowledge is related to the growth rate of the arrival of ideas in the following way:

$$1+g_A = \left(1+g_\alpha\right)^{\frac{1}{(1-\rho_l-\rho_m)}}.$$

As a result, the detrended equilibrium evolution of the local stock of knowledge evolves according to

$$\tilde{A}_{n,t+1} - \frac{\tilde{A}_{n,t}}{(1+g_A)} = \frac{\tilde{\alpha}_0\Gamma_\rho}{(1+g_A)} \sum_{i=1}^{N} s_{in,t}\left(\tilde{A}_{i,t}\right)^{\rho_l} \sum_{i=1}^{N} \lambda_{ni,t}\left(\frac{\tilde{A}_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m}$$

or

$$\frac{\tilde{A}_{n,t+1}}{\tilde{A}_{n,t}} = \frac{1}{1+g_A} + \frac{\alpha_0 \Gamma_\rho}{(1+g_A)\,\bar{A}_{n,t}} \sum_{i=1}^{N} s_{in,t} \left(\tilde{A}_{i,t}\right)^{\rho_l} \sum_{i=1}^{N} \lambda_{ni,t} \left(\frac{\tilde{A}_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m}.$$

We now consider the detrended value functions of the workers. Let $e^{V_{i,t}} = e^{\tilde{V}_{i,t}} (1+g_v)^t$. We then have

$$\tilde{V}_{i,t} + log\,(1{+}g_v)^t = log\left(\frac{\tilde{w}_{i,t}}{\tilde{P}_{i,t}}\left(1{+}g_{w/p}\right)^t\right) + \nu log\left(\sum_{n=1}^{N} exp\left(\beta \tilde{V}_{n,t+1} {+} \beta log\,(1{+}g_v)^{t+1} - m_{in,t}\right)^{1/\nu}\right), \quad \text{(C.1)}$$

where $g_{w/p}$ is the growth rate of $\tilde{w}_{i,t}/\tilde{P}_{i,t}$ at the balanced growth path. It follows that

$$\tilde{V}_{i,t} + log\,(1{+}g_v)^t = log\left(\frac{\tilde{w}_{i,t}}{\tilde{P}_{i,t}}\right) + log\left(1{+}g_{w/p}\right)^t + log\,(1{+}g_v)^{\beta(t+1)} + \nu log\left(\sum_{n=1}^{N} exp\left(\beta \tilde{V}_{n,t+1} - m_{in,t}\right)^{1/\nu}\right),$$

which immediately implies that

$$(1+g_v)^{(1-\beta)t} = \left(1+g_{w/p}\right)^t,$$

$$1+g_v = \left(1+g_{w/p}\right)^{\frac{1}{(1-\beta)}}.$$

Hence, the detrended equilibrium values become

$$\tilde{V}_{i,t} = log\left(\tilde{w}_{i,t}/\tilde{P}_{i,t}\right) + log\,(1+g_v)^\beta + \nu log\left(\sum_{n=1}^{N} exp\left(\beta \tilde{V}_{n,t+1} - m_{in,t}\right)^{1/\nu}\right).$$

Note that this result immediately implies that $\mu_{in,t}$ is not growing in the long run since

$$\begin{aligned}\mu_{in,t} &= \frac{exp\left(\beta V_{n,t+1} - m_{in,t}\right)^{1/\nu}}{\sum_{l=1}^{N} exp\left(\beta V_{l,t+1} - m_{il,t}\right)^{1/\nu}} \\ &= \frac{exp\left(\beta \tilde{V}_{n,t+1} - m_{in,t}\right)^{1/\nu}}{\sum_{l=1}^{N} exp\left(\beta \tilde{V}_{l,t+1} - m_{il,t}\right)^{1/\nu}}.\end{aligned}$$

It also implies that $L_{i,t}$ does not have long-run growth since

$$\begin{aligned}L_{i,t+1} &= \sum_{n=1}^{N} \mu_{ni,t} L_{n,t}, \\ &= \sum_{n=1}^{N} \frac{exp\left(\beta V_{i,t+1} - m_{ni,t}\right)^{1/\nu}}{\sum_{l=1}^{N} exp\left(\beta V_{l,t+1} - m_{nl,t}\right)^{1/\nu}} L_{n,t} \\ &= \sum_{n=1}^{N} \frac{exp\left(\beta \tilde{V}_{i,t+1} - m_{ni,t}\right)^{1/\nu}}{\sum_{l=1}^{N} exp\left(\beta \tilde{V}_{l,t+1} - m_{nl,t}\right)^{1/\nu}} L_{n,t}.\end{aligned}$$

Let us now consider the labor market clearing condition,

$$w_{i,t}L_{i,t} = \sum_{n=1}^{N} A_{i,t}\left(\frac{\kappa_{ni,t}x_{i,t}}{P_{n,t}/T}\right)^{-\theta} w_{n,t}L_{n,t}.$$

First note that

$$\begin{aligned} x_{i,t} = \tilde{x}_{i,t}\left(1+g_x\right)^t &= B\left(\left(\frac{\tilde{w}_{i,t}}{\tilde{P}_{i,t}}\left(1+g_{w/p}\right)^t\right)^{\xi}\left(\frac{\tilde{r}_{i,t}}{\tilde{P}_{i,t}}\left(1+g_{r/p}\right)^t\right)^{1-\xi}\right)^{\gamma}\tilde{P}_{i,t}\left(1+g_p\right)^t \\ &= \tilde{x}_{i,t}\left(1+g_{w/p}\right)^{t\xi\gamma}\left(1+g_{r/p}\right)^{t(1-\xi)\gamma}\left(1+g_p\right)^t. \end{aligned}$$

Using this expression, we can express the labor market clearing condition in a detrended form as

$$\tilde{w}_{i,t}\left(1{+}g_w\right)^t L_{i,t} = \sum_{n=1}^{N}\tilde{A}_{i,t}\left(1{+}g_A\right)^t\left(\frac{\kappa_{ni,t}\tilde{x}_{i,t}\left(1{+}g_{w/p}\right)^{t\xi\gamma}\left(1{+}g_{r/p}\right)^{t(1-\xi)\gamma}\left(1{+}g_p\right)^t}{\tilde{P}_{n,t}\left(1{+}g_p\right)^t/T}\right)^{-\theta}\tilde{w}_{n,t}\left(1{+}g_w\right)^t L_{n,t},$$

where we use the fact that $L_{i,t}$ does not grow in the long run. It follows that

$$1 = \left(1+g_A\right)^t\left(\left(1+g_{w/p}\right)^{t\xi\gamma}\left(1+g_{r/p}\right)^{t(1-\xi)\gamma}\right)^{-\theta},$$

$$\left(1+g_{w/p}\right)^{\theta\xi\gamma}\left(1+g_{r/p}\right)^{\theta(1-\xi)\gamma} = \left(1+g_A\right). \tag{C.2}$$

We follow the same steps for the capital accumulation equation. Then the detrended labor and capital market equilibrium conditions become

$$\tilde{w}_{i,t}L_{i,t} = \sum_{n=1}^{N}\tilde{A}_{i,t}\left(\frac{\kappa_{ni,t}\tilde{x}_{i,t}}{\tilde{P}_{n,t}/T}\right)^{-\theta}\tilde{w}_{n,t}L_{n,t},$$

$$\tilde{r}_{i,t}\tilde{K}_{i,t} = \sum_{n=1}^{N}\tilde{A}_{i,t}\left(\frac{\kappa_{ni,t}\tilde{x}_{i,t}}{\tilde{P}_{n,t}/T}\right)^{-\theta}\tilde{r}_{n,t}\tilde{K}_{n,t},$$

where $\tilde{K}_{n,t}$ is the detrended value of capital that we subsequently characterize.

We now detrend the price index equilibrium condition,

$$P_{i,t} = T\left(\sum_{n=1}^{N}A_{n,t}\left(\kappa_{in,t}x_{n,t}\right)^{-\theta}\right)^{-1/\theta},$$

which once detrended can be expressed as

$$\tilde{P}_{i,t}\left(1+g_p\right)^t = T\left(\sum_{n=1}^{N}\tilde{A}_{n,t}\left(1+g_A\right)^t\left(\kappa_{in,t}\tilde{x}_{n,t}\left(1+g_{w/p}\right)^{t\xi\gamma}\left(1+g_{r/p}\right)^{t(1-\xi)\gamma}\left(1+g_p\right)^t\right)^{-\theta}\right)^{-1/\theta}.$$

Using equation (C.2), we obtain the detrended equilibrium condition for the price index:

$$\tilde{P}_{i,t} = T \left(\sum_{n=1}^{N} \tilde{A}_{n,t} \left(\kappa_{in,t} \tilde{x}_{n,t} \right)^{-\theta} \right)^{-1/\theta}.$$

Now note that since in equilibrium we have that

$$w_{i,t} L_{i,t} = \frac{\xi}{1-\xi} r_{i,t} K_{i,t},$$

then

$$\frac{\tilde{w}_{i,t}}{\tilde{P}_{i,t}} \left(1+g_{w/p}\right)^t L_{i,t} = \frac{\xi}{1-\xi} \frac{\tilde{r}_{i,t}}{\tilde{P}_{i,t}} \left(1+g_{r/p}\right)^t \tilde{K}_{i,t} \left(1+g_k\right)^t,$$

which immediately implies that

$$1+g_{w/p} = \left(1+g_{r/p}\right)\left(1+g_k\right). \tag{C.3}$$

We now detrend the law of motion of capital accumulation,

$$K_{i,t+1} = \beta \left(r_{i,t}/P_{i,t} + (1-\delta) \right) K_{i,t},$$

which can be written as

$$\tilde{K}_{i,t+1} \left(1+g_k\right)^{t+1} = \beta \left(\left(1+g_{r/p}\right)^t \frac{\tilde{r}_{i,t}}{\tilde{P}_{i,t}} + (1-\delta) \right) \tilde{K}_{i,t} \left(1+g_k\right)^t$$

or

$$\tilde{K}_{i,t+1} = \frac{\beta}{(1+g_k)} \left(\left(1+g_{r/p}\right)^t \frac{\tilde{r}_{i,t}}{\tilde{P}_{i,t}} + (1-\delta) \right) \tilde{K}_{i,t}.$$

We then require that

$$g_{r/p} = 0, \Rightarrow, g_r = g_p,$$

and in this way, the detrended capital accumulation equation becomes

$$\tilde{K}_{i,t+1} = \frac{\beta}{(1+g_k)} \left(\frac{\tilde{r}_{i,t}}{\tilde{P}_{i,t}} + (1-\delta) \right) \tilde{K}_{i,t}.$$

From equation (C.2) we obtain that

$$1+g_{w/p} = \left(1+g_A\right)^{\frac{1}{\theta\xi\gamma}},$$

and from equation (C.2) we also obtain that

$$1+g_k = \left(1+g_A\right)^{\frac{1}{\theta\xi\gamma}}.$$

D Existence and Uniqueness

We now define the balanced growth path equilibrium conditions.

Equilibrium. *The equilibrium variables of the detrended model at the balanced growth path* $\left\{\bar{V}_i, \bar{w}_i, \bar{r}_i, \bar{P}_i, \bar{L}_i, \bar{K}_i, \bar{A}_i\right\}_{i=1}^{N}$ *solve the following system*

$$\bar{V}_i = log\left((1+g_v)^{\beta}\,\bar{w}_i/\bar{P}_i\right) + \nu log\left(\sum_{n=1}^{N} exp\left(\beta\bar{V}_n - \bar{m}_{in}\right)^{1/\nu}\right),$$

$$\bar{w}_i\bar{L}_i = \sum_{n=1}^{N}\bar{A}_i\left(\frac{\bar{\kappa}_{ni}\bar{x}_i}{\bar{P}_n/T}\right)^{-\theta}\bar{w}_n\bar{L}_n,$$

$$\bar{r}_i\bar{K}_i = \sum_{n=1}^{N}\bar{A}_i\left(\frac{\bar{\kappa}_{ni}\bar{x}_i}{\bar{P}_n/T}\right)^{-\theta}\bar{r}_n\bar{K}_n,$$

$$\bar{P}_i = T\left(\sum_{n=1}^{N}\bar{A}_n\left(\bar{\kappa}_{in}\bar{x}_n\right)^{-\theta}\right)^{-1/\theta},$$

$$\bar{L}_i = \sum_{n=1}^{N}\bar{\mu}_{ni}\bar{L}_n,$$

$$\bar{r}_i = \bar{P}_i\left(\left(1+g_k\right)/\beta - (1-\delta)\right),$$

$$\bar{A}_n = \frac{\alpha_0^n\Gamma_\rho}{g_A}\sum_{i=1}^{N}\bar{s}_{in}\left(\bar{A}_i\right)^{\rho_l}\sum_{i=1}^{N}\bar{\lambda}_{ni}\left(\frac{\bar{A}_i}{\bar{\lambda}_{ni}}\right)^{\rho_m},$$

with $\bar{x}_i = B\left(\bar{w}_i^{\xi}\bar{r}_i^{1-\xi}\right)^{\gamma}\bar{P}_i^{1-\gamma}$*, and shares given by* $\bar{\lambda}_{in} = \frac{\bar{A}_n(\bar{\kappa}_{in}\bar{x}_n)^{-\theta}}{(\bar{P}_i/\mathrm{T})^{-\theta}}$*,* $\bar{\mu}_{in} = \frac{exp(\beta\bar{V}_n - \bar{m}_{in})^{1/\nu}}{\sum_{h=1}^{N} exp(\beta\bar{V}_h - \bar{m}_{ih})^{1/\nu}}$*, and* $\bar{s}_{in} = \frac{\bar{\mu}_{in}\bar{L}_i}{\sum_{h=1}^{N}\bar{\mu}_{hn}\bar{L}_h}$.

We now prove the existence and uniqueness of the steady state equilibrium of the detrended model at the balanced growth path in different versions of the model. We start by rewriting these equilibrium conditions as follows.

From the first-order condition of the firm's problem, we have that $\bar{w}_i\bar{L}_i = \frac{\xi}{1-\xi}\bar{r}_i\bar{K}_i$.

It follows that $\bar{r}_i = \bar{w}_i\frac{\bar{L}_i(1-\xi)}{\bar{K}_i\xi}$ in the balanced growth path of the detrended model.

Hence, we have that

$$\frac{\bar{K}_i}{\bar{L}_i} = \frac{\bar{w}_i}{\bar{P}_i}\frac{(1-\xi)}{\xi}\frac{\beta}{(1+g_k)-\beta(1-\delta)}.$$

Now we use that $\bar{x}_i = B\left(\bar{w}_i^{\xi}\bar{r}_i^{1-\xi}\right)^{\gamma}\bar{P}_i^{1-\gamma}$ to obtain that

$$
\begin{aligned}
\bar{x}_i &= B\left(\bar{w}_i^{\xi}\bar{r}_i^{1-\xi}\right)^{\gamma}\bar{P}_i^{1-\gamma} = B\left(\bar{w}_i^{\xi}\left(\bar{w}_i\frac{\bar{L}_i\left(1-\xi\right)}{\bar{K}_i\xi}\right)^{1-\xi}\right)^{\gamma}\bar{P}_i^{1-\gamma} \\
&= B\left(\frac{\left(1-\xi\right)}{\xi}\right)^{\left(1-\xi\right)\gamma}\left(\bar{w}_i\right)^{\gamma}\left(\frac{\bar{L}_i}{\bar{K}_i}\right)^{\left(1-\xi\right)\gamma}\bar{P}_i^{1-\gamma} \\
&= B\left(\frac{\left(1-\xi\right)}{\xi}\right)^{\left(1-\xi\right)\gamma}\left(\frac{\left(\left(1+g_k\right)-\beta\left(1-\delta\right)\right)\xi}{\left(1-\xi\right)\beta}\right)^{\left(1-\xi\right)\gamma}\left(\frac{\bar{P}_i}{\bar{w}_i}\right)^{-\xi\gamma}\bar{P}_i \\
&= \Psi\left(\bar{w}_i\right)^{\xi\gamma}\left(\bar{P}_i\right)^{1-\xi\gamma},
\end{aligned}
$$

where $\Psi = B\left(\frac{(1-\xi)}{\xi}\right)^{(1-\xi)\gamma}\left(\frac{((1+g_k)-\beta(1-\delta))\xi}{(1-\xi)\beta}\right)^{(1-\xi)\gamma}$.

Hence, we can rewrite the labor market clearing condition as

$$
\begin{aligned}
\bar{w}_i\bar{L}_i &= \sum_{n=1}^{N}\bar{A}_i\left(\frac{\bar{\kappa}_{ni}\bar{x}_i}{\bar{P}_n/T}\right)^{-\theta}\bar{w}_n\bar{L}_n \\
&= \sum_{n=1}^{N}\bar{A}_i\left(\frac{\bar{\kappa}_{ni}\Psi\left(\bar{w}_i\right)^{\xi\gamma}\left(\bar{P}_i\right)^{1-\xi\gamma}}{\bar{P}_n/T}\right)^{-\theta}\bar{w}_n\bar{L}_n
\end{aligned}
$$

or

$$
\left(\bar{w}_i\right)^{1+\xi\gamma\theta}\bar{L}_i\left(\bar{P}_i\right)^{\theta(1-\xi\gamma)}\left(\bar{A}_i\right)^{-1} = \sum_{n=1}^{N}\left(\bar{\kappa}_{ni}\Psi T\right)^{-\theta}\left(\bar{P}_n\right)^{\theta}\bar{w}_n\bar{L}_n.
$$

Analogously, the price index can be written as

$$
\begin{aligned}
\bar{P}_i^{-\theta} &= T^{-\theta}\sum_{n=1}^{N}\bar{A}_n\left(\bar{\kappa}_{in}\bar{x}_n\right)^{-\theta} \\
&= T^{-\theta}\sum_{n=1}^{N}\bar{A}_n\bar{\kappa}_{in}^{-\theta}\Psi^{-\theta}\left(\bar{w}_n\right)^{-\xi\gamma\theta}\left(\bar{P}_n\right)^{-\theta(1-\xi\gamma)}.
\end{aligned}
$$

Turning to the value functions, we use the following change of variables:

$$
\begin{aligned}
\widetilde{m}_{in} &\equiv exp\left(\bar{m}_{in}\right)^{-1/\nu}, \\
\bar{\phi}_i &= \sum_{n=1}^{N}exp\left(\beta\bar{V}_n-\bar{m}_{in}\right)^{1/\nu}.
\end{aligned}
$$

Using these conditions, we express

$$
\exp\left(\frac{\beta}{\nu}\bar{V}_i\right) = \left(\zeta\bar{w}_i/\bar{P}_i\right)^{\frac{\beta}{\nu}}\left(\sum_{n=1}^{N}\widetilde{m}_{in}\exp\left(\frac{\beta}{\nu}\bar{V}_n\right)\right)^{\beta},
$$

$$\exp\left(\frac{\beta}{\nu}\bar{V}_i\right) = \left(\zeta\bar{w}_i/\bar{P}_i\right)^{\frac{\beta}{\nu}}\bar{\phi}_i^{\beta},$$

with $\zeta = (1+g_v)^{\beta}$. Hence,

$$\bar{\phi}_i = \sum_{n=1}^{N} \widetilde{m}_{in}\left(\zeta\bar{w}_n/\bar{P}_n\right)^{\frac{\beta}{\nu}}\bar{\phi}_n^{\beta}.$$

We also re-express the gross flows equation as

$$\bar{\mu}_{ni} = \frac{exp\left(\beta\bar{V}_i - \bar{m}_{ni}\right)^{1/\nu}}{\sum_{h=1}^{N} exp\left(\beta\bar{V}_h - \bar{m}_{nh}\right)^{1/\nu}} = \frac{\widetilde{m}_{ni}\left(\zeta\bar{w}_i/\bar{P}_i\right)^{\frac{\beta}{\nu}}\bar{\phi}_i^{\beta}}{\bar{\phi}_n}.$$

Hence, we express the law of motion of labor as

$$\begin{aligned}
\bar{L}_i &= \sum_{n=1}^{N}\bar{\mu}_{ni}\bar{L}_n,\\
\bar{L}_i &= \sum_{n=1}^{N}\frac{\widetilde{m}_{ni}\left(\zeta\bar{w}_i/\bar{P}_i\right)^{\frac{\beta}{\nu}}\bar{\phi}_i^{\beta}}{\bar{\phi}_n}\bar{L}_n,\\
\bar{w}_i^{-\frac{\beta}{\nu}}\bar{P}_i^{\frac{\beta}{\nu}}\bar{\phi}_i^{-\beta}\bar{L}_i \quad \zeta^{-\frac{\beta}{\nu}} &= \sum_{n=1}^{N}\widetilde{m}_{ni}\bar{\phi}_n^{-1}\bar{L}_n.
\end{aligned}$$

Finally, the evolution of technology is given by

$$\bar{A}_n = \frac{\alpha_0\Gamma_\rho}{g_A}\sum_{i=1}^{N}\bar{s}_{in}\left(\bar{A}_i\right)^{\rho_l}\sum_{i=1}^{N}\bar{\lambda}_{ni}\left(\frac{\bar{A}_i}{\bar{\lambda}_{ni}}\right)^{\rho_m}$$

and using $\bar{x}_i = B\left(\bar{w}_i^{\xi}\bar{r}_i^{1-\xi}\right)^{\gamma}\bar{P}_i^{1-\gamma}$, $\bar{\lambda}_{in} = \frac{\bar{A}_n(\bar{\kappa}_{in}\bar{x}_n)^{-\theta}}{(\bar{P}_i/T)^{-\theta}}$, $\bar{\mu}_{ni} = \frac{\widetilde{m}_{ni}(\zeta\bar{w}_i/\bar{P}_i)^{\frac{\beta}{\nu}}\bar{\phi}_i^{\beta}}{\bar{\phi}_n}$, and $\bar{s}_{in} = \frac{\bar{\mu}_{in}\bar{L}_i}{\sum_{h=1}^{N}\bar{\mu}_{hn}\bar{L}_h}$, we obtain

$$\begin{aligned}
\bar{A}_n &= \frac{\alpha_0\Gamma_\rho}{g_A}\sum_{i=1}^{N}\bar{s}_{in}\left(\bar{A}_i\right)^{\rho_l}\sum_{i=1}^{N}\left(\bar{\lambda}_{ni}\right)^{1-\rho_m}\left(\bar{A}_i\right)^{\rho_m}\\
&=\frac{\alpha_0\Gamma_\rho}{g_A}\sum_{i=1}^{N}\frac{\bar{\mu}_{in}\bar{L}_i}{\bar{L}_n}\left(\bar{A}_i\right)^{\rho_l}\sum_{i=1}^{N}\left(\frac{\left(\bar{\kappa}_{ni}\bar{x}_i\right)^{-\theta}}{\left(\bar{P}_n/T\right)^{-\theta}}\right)^{1-\rho_m}\bar{A}_i\\
&=\frac{\alpha_0\Gamma_\rho}{g_A}\sum_{i=1}^{N}\frac{L_i}{L_n}\frac{\widetilde{m}_{in}\left(\zeta\bar{w}_n/\bar{P}_n\right)^{\frac{\beta}{\nu}}\bar{\phi}_n^{\beta}}{\bar{\phi}_i}\left(\bar{A}_i\right)^{\rho_l}\sum_{i=1}^{N}\left(\frac{\left(\bar{\kappa}_{ni}\Psi\left(\bar{w}_i\right)^{\xi\gamma}\left(\bar{P}_i\right)^{1-\xi\gamma}\right)^{-\theta}}{\left(\bar{P}_n/T\right)^{-\theta}}\right)^{1-\rho_m}\bar{A}_i,
\end{aligned}$$

and hence,

$$\bar{A}_n\bar{P}_n^{\frac{\beta}{\nu}-\theta(1-\rho_m)}\bar{w}_n^{-\frac{\beta}{\nu}}\bar{\phi}_n^{-\beta}\bar{L}_n = \frac{\alpha_t^n\Gamma_\rho}{\zeta^{-\frac{\beta}{\nu}}g_A}\sum_{i=1}^{N}\widetilde{m}_{in}L_i\bar{\phi}_i^{-1}\bar{A}_i^{\rho_l}\sum_{i=1}^{N}\left(\bar{\kappa}_{ni}\Psi T\right)^{-\theta(1-\rho_m)}\bar{w}_i^{-\xi\gamma\theta(1-\rho_m)}\bar{P}_i^{-(1-\xi\gamma)\theta(1-\rho_m)}\bar{A}_i.$$

Finally, rearranging this expression, we obtain

$$\bar{w}_n^{-\frac{\beta}{\nu}} \bar{P}_n^{\frac{\beta}{\nu}-(1-\rho_m)\theta} \bar{\phi}_n^{-\beta} \bar{A}_n \bar{L}_n = \frac{\alpha_0 \Gamma_\rho \left(\Psi T\right)^{-(1-\rho_m)\theta}}{\zeta^{-\frac{\beta}{\nu}} g_A} \sum_{i=1}^{N} \widetilde{m}_{in} \bar{\phi}_i^{-1} \bar{A}_i^{\rho_l} \bar{L}_i \sum_{i=1}^{N} \bar{\kappa}_{ni}^{-(1-\rho_m)\theta} \bar{w}_i^{-(1-\rho_m)\theta\xi\gamma} \bar{P}_i^{-(1-\rho_m)\theta(1-\xi\gamma)} \bar{A}_i.$$

We end up with the following system of equations to solve for the equilibrium variables at the balanced growth path:

$$\bar{w}_i^{1+\xi\gamma\theta} \bar{P}_i^{\theta(1-\xi\gamma)} \bar{L}_i \bar{A}_i^{-1} = \left(\Psi T\right)^{-\theta} \sum_{n=1}^{N} \bar{\kappa}_{ni}^{-\theta} \bar{w}_n \bar{P}_n^{\theta} \bar{L}_n, \tag{D.1}$$

$$\bar{P}_i^{-\theta} = \left(\Psi T\right)^{-\theta} \sum_{n=1}^{N} \bar{\kappa}_{in}^{-\theta} \bar{w}_n^{-\theta\xi\gamma} \bar{P}_n^{-\theta(1-\xi\gamma)} \bar{A}_n, \tag{D.2}$$

$$\bar{w}_i^{-\frac{\beta}{\nu}} \bar{P}_i^{\frac{\beta}{\nu}} \bar{L}_i \bar{\phi}_i^{-\beta} \zeta^{-\frac{\beta}{\nu}} = \sum_{n=1}^{N} \widetilde{m}_{ni} \bar{L}_n \bar{\phi}_n^{-1}, \tag{D.3}$$

$$\bar{\phi}_i = \sum_{n=1}^{N} \widetilde{m}_{in} \zeta^{\frac{\beta}{\nu}} \bar{w}_n^{\frac{\beta}{\nu}} \bar{P}_n^{-\frac{\beta}{\nu}} \bar{\phi}_n^{\beta}, \tag{D.4}$$

$$\bar{w}_n^{-\frac{\beta}{\nu}} \bar{P}_n^{\frac{\beta}{\nu}-(1-\rho_m)\theta} \bar{\phi}_n^{-\beta} \bar{A}_n \bar{L}_n = \frac{\zeta^{\frac{\beta}{\nu}} \alpha_0 \Gamma_\rho}{g_A (\Psi T)^{(1-\rho_m)\theta}} \sum_{i=1}^{N} \widetilde{m}_{in} \bar{\phi}_i^{-1} \bar{A}_i^{\rho_l} \bar{L}_i \sum_{i=1}^{N} \kappa_{ni}^{-(1-\rho_m)\theta} \bar{w}_i^{-(1-\rho_m)\theta\xi\gamma} \bar{P}_i^{-(1-\rho_m)\theta(1-\xi\gamma)} \bar{A}_i. \tag{D.5}$$

We now proceed to prove the existence and uniqueness of the steady state equilibrium of the detrended model at the balanced growth path in different versions of the model.

D.0.1 Case 1: Model with No Idea Flows

Without idea flows, the steady state equilibrium is characterized by the following set of equilibrium conditions:

$$\bar{w}_i^{1+\xi\gamma\theta} \bar{P}_i^{\theta(1-\xi\gamma)} \bar{L}_i \bar{A}_i^{-1} = \left(\Psi T\right)^{-\theta} \sum_{n=1}^{N} \bar{\kappa}_{ni}^{-\theta} \bar{w}_n \bar{P}_n^{\theta} \bar{L}_n, \tag{D.6}$$

$$\bar{P}_i^{-\theta} = \left(\Psi T\right)^{-\theta} \sum_{n=1}^{N} \bar{\kappa}_{in}^{-\theta} \bar{w}_n^{-\theta\xi\gamma} \bar{P}_n^{-\theta(1-\xi\gamma)} \bar{A}_n, \tag{D.7}$$

$$\bar{w}_i^{-\frac{\beta}{\nu}} \bar{P}_i^{\frac{\beta}{\nu}} \bar{L}_i \bar{\phi}_i^{-\beta} = \sum_{n=1}^{N} \widetilde{m}_{ni} \bar{L}_n \bar{\phi}_n^{-1}, \tag{D.8}$$

$$\bar{\phi}_i = \sum_{n=1}^{N} \widetilde{m}_{in} \bar{w}_n^{\frac{\beta}{\nu}} \bar{P}_n^{-\frac{\beta}{\nu}} \bar{\phi}_n^{\beta}. \tag{D.9}$$

We can write the matrices Λ and Γ representing the exponents of $\{\bar{w}_i, \bar{P}_i, \bar{L}_i, \bar{\phi}_i\}$ on the left-hand side and right-hand side of the system of equations, respectively. These matrices are given by

$$\Lambda=\left(\begin{array}{cccc} 1+\theta\xi\gamma & \theta\left(1-\xi\gamma\right) & 1 & 0\\ 0 & -\theta & 0 & 0\\ -\frac{\beta}{\nu} & \frac{\beta}{\nu} & 1 & -\beta\\ 0 & 0 & 0 & 1 \end{array}\right),\ \Gamma=\left(\begin{array}{cccc} 1 & \theta & 1 & 0\\ -\theta\xi\gamma & -\theta\left(1-\xi\gamma\right) & 0 & 0\\ 0 & 0 & 1 & -1\\ \dfrac{\beta}{\nu} & -\dfrac{\beta}{\nu} & 0 & \beta \end{array}\right).$$

We then define the matrix $\Omega=\Gamma\Lambda^{-1}$. Following Kleinman, Liu, and Redding (2021) and Allen, Arkolakis, and Li (2020), we show that if the spectral radius of Ω is equal to one ($\rho\left(\Omega\right)=1$) and if Ω is invertible, then the solution must be unique up to scale. Evaluating the eigenvalues of Ω, we have

$$\left[\begin{array}{c} \zeta_1\\ \zeta_2\\ \zeta_3\\ \zeta_4 \end{array}\right]=\left[\begin{array}{c} 1\\ 1\\ \frac{-b+\sqrt{b^2-4ac}}{2a}\\ \frac{-b-\sqrt{b^2-4ac}}{2a} \end{array}\right],$$

where

$$\begin{aligned} a &= \beta+\nu+\theta\gamma\nu\xi, && \text{(D.10)}\\ b &= -\nu\left(1+\beta-\gamma\xi\left(1+\theta\left(1-\beta\right)\right)\right), && \text{(D.11)}\\ c &= \beta\left(\nu-1-\gamma\xi\nu\left(1+\theta\right)\right). && \text{(D.12)} \end{aligned}$$

First, $\left|-b+\sqrt{b^2-4ac}\right|<2a$.

If $-b+\sqrt{b^2-4ac}>0$, then this is equivalent to show that $b^2-4ac<\left(2a+b\right)^2$ or that $b^2-4ac<\left(2a+b\right)^2=4a^2+4ab+b^2$, or equivalently that $0<a+b+c$ which holds since $a+b+c=\gamma\nu\xi(1-\beta)(1+2\theta)>0$.

Otherwise, if $-b+\sqrt{b^2-4ac}<0$, we need to show that $b-\sqrt{b^2-4ac}<2a$, or that $-\sqrt{b^2-4ac}<2a-b$ but note that $2a-b=2\left(\beta+\nu+\theta\gamma\nu\xi\right)+\nu\left(1+\beta-\gamma\xi\left(1+\theta\left(1-\beta\right)\right)\right)=2\beta+3\nu+\beta\nu+\left(\theta-1\right)\gamma\nu\xi+\theta\beta\gamma\nu\xi>0$ so it holds since $\theta\geq 0$. Note that at $\theta=0$, $2\beta+3\nu+\beta\nu-\gamma\nu\xi>0$.

Second, $\left|-b-\sqrt{b^2-4ac}\right|<2a$.

If $-b-\sqrt{b^2-4ac}<0$ then this is equivalent to show that $b+\sqrt{b^2-4ac}<2a$, or that $\sqrt{b^2-4ac}<2a-b$, and given that $2a-b=2\left(\beta+\nu+\theta\gamma\nu\xi\right)+\nu\left(1+\beta-\gamma\xi\left(1+\theta\left(1-\beta\right)\right)\right)=2\beta+3\nu+\beta\nu+\left(\theta-1\right)\gamma\xi\nu+\theta\beta\gamma\nu\xi>0$ then this is equivalent to show that $b^2-4ac<\left(2a-b\right)^2$, $b^2-4ac<\left(2a-b\right)^2=4a^2+b^2-4ab$ or $-c<a-b$, or $0<a-b+c=\nu(1+\beta)(2-\gamma\xi)>0$.

If $-b-\sqrt{b^2-4ac}>0$ then this is equivalent to show that $-b-\sqrt{b^2-4ac}<2a$, or that $0<2a+b$, but since we know that $a+b=\beta(1-\nu)+\gamma\nu\xi+(2-\beta)\theta\gamma\nu\xi>0$, then $0<2a+b$. *Q.E.D.*

D.0.2 Case 2: Model with Idea Flows from Migration

With idea flows from migration, the balanced growth path in the detrended equilibrium is characterized by the following set of equilibrium conditions:

$$\bar{w}_i^{1+\xi\gamma\theta}\bar{P}_i^{\theta(1-\xi\gamma)}\bar{L}_i\bar{A}_i^{-1} = (\Psi T)^{-\theta}\sum_{n=1}^{N}\bar{\kappa}_{ni}^{-\theta}\bar{w}_n\bar{P}_n^{\theta}\bar{L}_n, \tag{D.13}$$

$$\bar{P}_i^{-\theta} = (\Psi T)^{-\theta}\sum_{n=1}^{N}\bar{\kappa}_{in}^{-\theta}\bar{w}_n^{-\theta\xi\gamma}\bar{P}_n^{-\theta(1-\xi\gamma)}\bar{A}_n, \tag{D.14}$$

$$\bar{w}_i^{-\frac{\beta}{\nu}}\bar{P}_i^{\frac{\beta}{\nu}}\bar{L}_i\bar{\phi}_i^{-\beta}\zeta^{-\frac{\beta}{\nu}} = \sum_{n=1}^{N}\widetilde{m}_{ni}\bar{L}_n\bar{\phi}_n^{-1}, \tag{D.15}$$

$$\bar{\phi}_i = \sum_{n=1}^{N}\widetilde{m}_{in}\zeta^{\frac{\beta}{\nu}}\bar{w}_n^{\frac{\beta}{\nu}}\bar{P}_n^{-\frac{\beta}{\nu}}\bar{\phi}_n^{\beta}, \tag{D.16}$$

$$\bar{w}_i^{-\frac{\beta}{\nu}}\bar{P}_i^{\frac{\beta}{\nu}}\bar{L}_i\bar{\phi}_i^{-\beta}\bar{A}_i\zeta^{-\frac{\beta}{\nu}} = \frac{\alpha_0\Gamma_\rho}{g_A}\sum_{n=1}^{N}\widetilde{m}_{ni}\bar{L}_n\bar{\phi}_n^{-1}\left(\bar{A}_n\right)^{\rho_l}. \tag{D.17}$$

Analogous to the previous case, we can write the matrices Λ and Γ representing the exponents of $\{\bar{w}_i, \bar{P}_i, \bar{L}_i, \bar{\phi}_i, \bar{A}_i\}$ on the left-hand side and right-hand side of the system of equations, respectively. These matrices are given by

$$\Lambda = \begin{pmatrix} 1+\theta\xi\gamma & \theta(1-\xi\gamma) & 1 & 0 & -1 \\ 0 & -\theta & 0 & 0 & 0 \\ -\frac{\beta}{\nu} & \frac{\beta}{\nu} & 1 & -\beta & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -\frac{\beta}{\nu} & \frac{\beta}{\nu} & 1 & -\beta & 1 \end{pmatrix},$$

$$\Gamma = \begin{pmatrix} 1 & \theta & 1 & 0 & 0 \\ -\theta\xi\gamma & -\theta(1-\xi\gamma) & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 \\ \dfrac{\beta}{\nu} & -\dfrac{\beta}{\nu} & 0 & \beta & 0 \\ 0 & 0 & 1 & -1 & \rho_l \end{pmatrix}.$$

As in the previous case, evaluating the eigenvalues of Ω, we have

$$\begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \frac{-b+\sqrt{b^2-4ac}}{2a} \\ \frac{-b-\sqrt{b^2-4ac}}{2a} \\ \rho_l \end{bmatrix},$$

where

$$\begin{aligned}
a &= \beta+\nu+\theta\gamma\nu\xi, \\
b &= -\nu\left(1+\beta-\gamma\xi\left(1+\theta\left(1-\beta\right)\right)\right), \\
c &= \beta\left(\nu-1-\gamma\xi\nu\left(1+\theta\right)\right).
\end{aligned}$$

Hence, the first four eigenvalues are the same as in the previous case, and the additional eigenvalue is given by $\rho_l < 1$. It follows that the balanced growth path equilibrium of the detrended model with idea flows from migration is unique up to scale.

D.0.3 Case 3: Model with Idea Flows from Trade

With idea flows from trade, the equilibrium is characterized by the following set of equilibrium conditions:

$$\bar{w}_i^{1+\xi\gamma\theta}\bar{P}_i^{\theta(1-\xi\gamma)}\bar{L}_i\bar{A}_i^{-1} = (\Psi T)^{-\theta}\sum_{n=1}^{N}\bar{\kappa}_{ni}^{-\theta}\bar{w}_n\bar{P}_n^{\theta}\bar{L}_n, \tag{D.18}$$

$$\bar{P}_i^{-\theta} = (\Psi T)^{-\theta}\sum_{n=1}^{N}\bar{\kappa}_{in}^{-\theta}\bar{w}_n^{-\theta\xi\gamma}\bar{P}_n^{-\theta(1-\xi\gamma)}\bar{A}_n, \tag{D.19}$$

$$\bar{w}_i^{-\frac{\beta}{\nu}}\bar{P}_i^{\frac{\beta}{\nu}}\bar{L}_i\bar{\phi}_i^{-\beta}\zeta^{-\frac{\beta}{\nu}} = \sum_{n=1}^{N}\widetilde{m}_{ni}\bar{L}_n\bar{\phi}_n^{-1}, \tag{D.20}$$

$$\bar{\phi}_i = \sum_{n=1}^{N}\ \widetilde{m}_{in}\zeta^{\frac{\beta}{\nu}}\bar{w}_n^{\frac{\beta}{\nu}}\bar{P}_n^{-\frac{\beta}{\nu}}\bar{\phi}_n^{\beta}, \tag{D.21}$$

$$\bar{P}_i^{-(1-\rho_m)\theta}\bar{A}_i = \frac{\alpha_0\Gamma_\rho\left(T\Psi\right)^{-(1-\rho_m)\theta}}{g_A}\sum_{n=1}^{N}\bar{\kappa}_{in}^{-(1-\rho_m)\theta}\bar{w}_n^{-(1-\rho_m)\theta\xi\gamma}\bar{P}_n^{-(1-\rho_m)\theta(1-\xi\gamma)}\bar{A}_n. \tag{D.22}$$

Analogous to the case of idea flows from migration, we can write the matrices Λ and Γ representing the exponents of $\{\bar{w}_i, \bar{P}_i, \bar{L}_i, \bar{\phi}_i, \bar{A}_i\}$ on the left-hand side and right-hand side of the system of equations, respectively. These matrices are given by

$$\Lambda = \begin{pmatrix} 1+\theta\xi\gamma & \theta\left(1-\xi\gamma\right) & 1 & 0 & -1 \\ 0 & -\theta & 0 & 0 & 0 \\ -\frac{\beta}{\nu} & \frac{\beta}{\nu} & 1 & -\beta & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -\left(1-\rho_m\right)\theta & 0 & 0 & 1 \end{pmatrix},$$

$$
\Gamma = \begin{pmatrix} 1 & \theta & 1 & 0 & 0 \\ -\theta\xi\gamma & -\theta\left(1-\xi\gamma\right) & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 \\ \frac{\beta}{\nu} & -\frac{\beta}{\nu} & 0 & \beta & 0 \\ -\left(1-\rho_m\right)\theta\xi\gamma & -\left(1-\rho_m\right)\theta\left(1-\xi\gamma\right) & 0 & 0 & 1 \end{pmatrix}.
$$

As before, evaluating the eigenvalues of Ω, we have

$$
\begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \frac{-b+\sqrt{b^2-4ac}}{2a} \\ \frac{-b-\sqrt{b^2-4ac}}{2a} \\ \rho_m \end{bmatrix},
$$

where

$$
\begin{aligned}
a &= \beta+\nu+\theta\gamma\nu\xi, \\
b &= -\nu\left(1+\beta-\gamma\xi\left(1+\theta\left(1-\beta\right)\right)\right), \\
c &= \beta\left(\nu-1-\gamma\xi\nu\left(1+\theta\right)\right).
\end{aligned}
$$

Hence, the five eigenvalues are the same as in the case with no idea flows (Case 1) except that now the fifth eigenvalue is given by $\rho_{m.}$ Since $\rho_m < 1$, it follows that the balanced growth path of the detrended equilibrium with idea flows from trade is unique up to scale.

D.0.4 Case 4: Model with Idea Flows from Migration and Trade

The balanced growth path equilibrium conditions of the detrended model were derived previously in this appendix.

To prove the existence and uniqueness of the balanced growth path equilibrium in the detrended model, we start with the following change of variables:

$$
\bar{A}_n = \bar{A}_n^l \bar{A}_n^m,
$$

$$
\bar{A}_n^l = \sum_{i=1}^{N} \frac{\bar{\mu}_{in}\bar{L}_i}{\bar{L}_n}\left(\bar{A}_i\right)^{\rho_l},
$$

$$
\bar{A}_n^m = \frac{\alpha_0\Gamma_\rho}{g_A}\sum_{i=1}^{N}\bar{\lambda}_{ni}\left(\frac{\bar{A}_i}{\bar{\lambda}_{ni}}\right)^{\rho_m}.
$$

Using these expressions in the equilibrium derived at the beginning of this appendix, we obtain the equilibrium conditions with idea flows from migration and trade:

$$\bar{w}_i^{1+\xi\gamma\theta}\bar{P}_i^{\theta(1-\xi\gamma)}\bar{L}_i\left(\bar{A}_i^l\bar{A}_i^m\right)^{-1}=(\Psi T)^{-\theta}\sum_{n=1}^{N}\bar{\kappa}_{ni}^{-\theta}\bar{w}_n\bar{P}_n^{\theta}\bar{L}_n, \tag{D.23}$$

$$\bar{P}_i^{-\theta}=(\Psi T)^{-\theta}\sum_{n=1}^{N}\bar{\kappa}_{in}^{-\theta}\bar{w}_n^{-\theta\xi\gamma}\bar{P}_n^{-\theta(1-\xi\gamma)}\bar{A}_n^l\bar{A}_n^m, \tag{D.24}$$

$$\bar{w}_i^{-\frac{\beta}{\nu}}\bar{P}_i^{\frac{\beta}{\nu}}\bar{L}_i\bar{\phi}_i^{-\beta}\zeta^{-\frac{\beta}{\nu}}=\sum_{n=1}^{N}\widetilde{m}_{ni}\bar{L}_n\bar{\phi}_n^{-1}, \tag{D.25}$$

$$\bar{\phi}_i=\sum_{n=1}^{N}\ \widetilde{m}_{in}\zeta^{\frac{\beta}{\nu}}\bar{w}_n^{\frac{\beta}{\nu}}\bar{P}_n^{-\frac{\beta}{\nu}}\bar{\phi}_n^{\beta}, \tag{D.26}$$

$$\bar{w}_i^{-\frac{\beta}{\nu}}\bar{P}_i^{\frac{\beta}{\nu}}\bar{\phi}_i^{-\beta}\bar{A}_i^l\bar{L}_i\zeta^{-\frac{\beta}{\nu}}=\sum_{n=1}^{N}\widetilde{m}_{ni}\bar{\phi}_n^{-1}\left(\bar{A}_n^l\bar{A}_n^m\right)^{\rho_l}\bar{L}_n, \tag{D.27}$$

$$\bar{P}_i^{-(1-\rho_m)\theta}\bar{A}_i^m=\varpi\sum_{n=1}^{N}\kappa_{in}^{-(1-\rho_m)\theta}\bar{w}_n^{-(1-\rho_m)\theta\xi\gamma}\bar{P}_n^{-(1-\rho_m)\theta(1-\xi\gamma)}\bar{A}_n^l\bar{A}_n^m, \tag{D.28}$$

where $\varpi=\frac{\alpha_0\Gamma_\rho(\Psi T)^{-(1-\rho_m)\theta}}{g_A}$. Analogous to the previous cases, we can write the matrices Λ and Γ representing the exponents of $\{\bar{w}_i,\bar{P}_i,\bar{L}_i,\bar{\phi}_i,\bar{A}_i^l,\bar{A}_i^m\}$ on the left-hand side and right-hand side of the system of equations, respectively. These matrices are given by

$$\Lambda=\begin{pmatrix} 1+\theta\xi\gamma & \theta(1-\xi\gamma) & 1 & 0 & -1 & -1\\ 0 & -\theta & 0 & 0 & 0 & 0\\ -\frac{\beta}{\nu} & \frac{\beta}{\nu} & 1 & -\beta & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ -\frac{\beta}{\nu} & \frac{\beta}{\nu} & 1 & -\beta & 1 & 0\\ 0 & -(1-\rho_m)\theta & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\Gamma=\begin{pmatrix} 1 & \theta & 1 & 0 & 0 & 0\\ -\theta\xi\gamma & -\theta(1-\xi\gamma) & 0 & 0 & 1 & 1\\ 0 & 0 & 1 & -1 & 0 & 0\\ \dfrac{\beta}{\nu} & -\dfrac{\beta}{\nu} & 0 & \beta & 0 & 0\\ 0 & 0 & 1 & -1 & \rho_l & \rho_l\\ -(1-\rho_m)\theta\xi\gamma & -(1-\rho_m)\theta(1-\xi\gamma) & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Evaluating the eigenvalues of Ω, we have

$$\begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \\ \zeta_6 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \frac{-b+\sqrt{b^2-4ac}}{2a} \\ \frac{-b-\sqrt{b^2-4ac}}{2a} \\ 0 \\ \rho_m + \rho_l \end{bmatrix},$$

where

$$\begin{aligned} a &= \beta + \nu + \theta\gamma\nu\xi, \\ b &= -\nu\left(1 + \beta - \gamma\xi\left(1 + \theta\left(1 - \beta\right)\right)\right), \\ c &= \beta\left(\nu - 1 - \gamma\xi\nu\left(1 + \theta\right)\right). \end{aligned}$$

Hence, the first four eigenvalues are the same as in the previous cases. The two additional eigenvalues are zero and $\rho_m + \rho_l < 1$. It follows that the balanced growth path of the detrended equilibrium with idea flows from migration and trade is unique up to scale.

E Models in Changes

In this section of the appendix, we show how to express the equilibrium conditions of the dynamic spatial growth model in relative time differences.

E.1 Equilibrium Conditions in Relative Time Differences

Let us define $\hat{y}_{t+1}$ as the time difference in the detrended variable $\tilde{y}$; namely, $\hat{y}_{t+1} = (\tilde{y}_{t+1}/\tilde{y}_t)$. The equilibrium conditions in time differences of the detrended system are given by

$$\log(\hat{u}_{i,t+1}) = \log\left(\hat{w}_{i,t+1}/\hat{P}_{i,t+1}\right) + \nu \log\left(\sum_{n=1}^{N} \mu_{in,t}\left(\hat{u}_{n,t+2}\right)^{\beta/\nu}\left(\hat{m}_{in,t+1}\right)^{-1/\nu}\right), \tag{E.1}$$

$$\mu_{in,t+1} = \frac{\mu_{in,t}\left(\hat{u}_{n,t+2}\right)^{\beta/\nu}\left(\hat{m}_{in,t+1}\right)^{-1/\nu}}{\sum_{h=1}^{N}\mu_{ih,t}\left(\hat{u}_{h,t+2}\right)^{\beta/\nu}\left(\hat{m}_{ih,t+1}\right)^{-1/\nu}}, \tag{E.2}$$

$$L_{i,t+1} = \sum_{n=1}^{N} \mu_{ni,t} L_{n,t}, \tag{E.3}$$

$$\hat{x}_{i,t} = \left(\hat{w}_{i,t}^{\xi}\hat{r}_{i,t}^{1-\xi}\right)^{\gamma}\hat{P}_{i,t}^{1-\gamma}, \tag{E.4}$$

$$\hat{P}_{i,t+1} = \left(\sum_{n=1}^{N} \lambda_{in,t} \hat{A}_{n,t+1} \left(\hat{\kappa}_{in,t+1} \hat{x}_{n,t+1} \right)^{-\theta} \right)^{-1/\theta}, \tag{E.5}$$

$$\lambda_{in,t+1} = \lambda_{in,t} \hat{A}_{n,t+1} \left(\frac{\hat{\kappa}_{in,t+1} \hat{x}_{n,t+1}}{\hat{P}_{i,t+1}} \right)^{-\theta}, \tag{E.6}$$

$$\hat{w}_{i,t+1} \hat{L}_{i,t+1} = \frac{1}{\tilde{w}_{i,t} L_{i,t}} \sum_{n=1}^{N} \lambda_{ni,t+1} \hat{w}_{n,t+1} \hat{L}_{n,t+1} \tilde{w}_{n,t} L_{n,t}, \tag{E.7}$$

$$\tilde{K}_{i,t+1} = \frac{\beta}{(1+g_k)} \tilde{R}_{i,t} \tilde{K}_{i,t}, \tag{E.8}$$

$$\tilde{R}_{i,t+1} = 1 - \delta + \frac{\hat{w}_{i,t+1} \hat{L}_{i,t+1}}{\hat{P}_{i,t+1} \hat{K}_{i,t+1}} \left[\tilde{R}_{i,t} - (1-\delta) \right], \tag{E.9}$$

$$\hat{A}_{n,t+1} = \frac{1}{(1+g_A)} + \frac{\alpha_0 \Gamma_\rho}{\tilde{A}_{n,t}(1+g_A)} \sum_{i=1}^{N} s_{in,t} \left(\tilde{A}_{i,t} \right)^{\rho_\ell} \left[\sum_{i=1}^{N} \lambda_{ni,t} \left(\frac{\tilde{A}_{i,t}}{\lambda_{ni,t}} \right)^{\rho_m} \right], \tag{E.10}$$

where $\hat{u}_{i,t+1} = \exp(\tilde{V}_{i,t+1} - \tilde{V}_{i,t})$, $\hat{m}_{in,t+1} = \exp(m_{in,t+1} - m_{in,t})$, $\tilde{R}_{i,t} = \tilde{r}_{i,t} / \tilde{P}_{i,t} + (1-\delta)$. Note we use the fact that $L_{n,t} = \tilde{L}_{n,t}, \mu_{ni,t} = \tilde{\mu}_{ni,t}$ and $\lambda_{in,t} = \tilde{\lambda}_{in,t}$. We now turn to derive these equilibrium conditions.

To derive the system of equations in time differences, recall first that the share of workers moving from location i to n at time $t+1$ is given by

$$\begin{aligned}
\mu_{in,t+1} &= \frac{exp\left(\beta V_{n,t+2} - m_{in,t+1}\right)^{1/\nu}}{\sum_{h=1}^{N} exp\left(\beta V_{h,t+2} - m_{ih,t+1}\right)^{1/\nu}}, \\
&= \frac{exp\left(\beta \tilde{V}_{n,t+2} - m_{in,t+1}\right)^{1/\nu}}{\sum_{h=1}^{N} exp\left(\beta \tilde{V}_{h,t+2} - m_{ih,t+1}\right)^{1/\nu}},
\end{aligned}$$

where for the second equality we use the definition $e^{V_{i,t}} = e^{\tilde{V}_{i,t}} (1+g_v)^t$ for all i and t.

By multiplying and dividing $\mu_{in,t}$ in the numerator and $\mu_{ih,t}$ for each term in the summation in the denominator, we have

$$\begin{aligned}
\mu_{in,t+1} &= \frac{\mu_{in,t} exp\left(\beta \tilde{V}_{n,t+2} - \beta \tilde{V}_{n,t+1} + m_{in,t+1} - m_{in,t}\right)^{1/\nu}}{\sum_{h=1}^{N} \mu_{ih,t} exp\left(\beta \tilde{V}_{h,t+2} - \beta \tilde{V}_{h,t+1} + m_{ih,t+1} - m_{ih,t}\right)^{1/\nu}} \\
&= \frac{\mu_{in,t} \left(\hat{u}_{n,t+2}\right)^{\beta/\nu} \left(\hat{m}_{in,t+1}\right)^{-1/\nu}}{\sum_{h=1}^{N} \mu_{ih,t} \left(\hat{u}_{h,t+2}\right)^{\beta/\nu} \left(\hat{m}_{ih,t+1}\right)^{-1/\nu}},
\end{aligned}$$

which is equation (E.2).

To obtain equilibrium condition (E.1), we take the time difference using (10). We obtain

$$
\begin{aligned}
\log(\hat{u}_{i,t+1}) &= \tilde{V}_{i,t+1} - \tilde{V}_{i,t} \\
&= \log\left(\frac{\tilde{w}_{i,t+1}/\tilde{P}_{i,t+1}}{\tilde{w}_{i,t}/\tilde{P}_{i,t}}\right) + \nu \log\left(\sum_{n=1}^{N} \frac{\exp\left(\beta \tilde{V}_{n,t+2} - m_{in,t+1}\right)^{1/\nu}}{\sum_{h=1}^{N} \exp\left(\beta \tilde{V}_{h,t+1} - m_{ih,t}\right)^{1/\nu}}\right) \\
&= \log\left(\hat{w}_{i,t+1}/\hat{P}_{i,t+1}\right) + \nu \log\left(\sum_{n=1}^{N} \frac{\exp\left(\beta \tilde{V}_{n,t+1} - m_{in,t}\right)^{1/\nu} \frac{\exp\left(\beta \tilde{V}_{n,t+2} - m_{in,t+1}\right)^{1/\nu}}{\exp\left(\beta \tilde{V}_{n,t+1} - m_{in,t}\right)^{1/\nu}}}{\sum_{h=1}^{N} \exp\left(\beta \tilde{V}_{h,t+1} - m_{ih,t}\right)^{1/\nu}}\right) \\
&= \log\left(\hat{w}_{i,t+1}/\hat{P}_{i,t+1}\right) + \nu \log\left(\sum_{n=1}^{N} \mu_{in,t} \exp\left(\tilde{V}_{n,t+2} - \tilde{V}_{n,t+1}\right)^{\beta/\nu} \exp\left(m_{in,t+1} - m_{in,t}\right)^{-1/\nu}\right) \\
&= \log\left(\hat{w}_{i,t+1}/\hat{P}_{i,t+1}\right) + \nu \log\left(\sum_{n=1}^{N} \mu_{in,t} \left(\hat{u}_{n,t+2}\right)^{\beta/\nu} \left(\hat{m}_{in,t+1}\right)^{-1/\nu}\right),
\end{aligned}
$$

where for the third equality we use the expression of $\mu_{in,t}$ previously derived.

Since labor in each location is constant in the long run, we immediately obtain (E.3) from the law of motion (14).

To obtain equation (E.4), note that

$$
\tilde{x}_{i,t} = B\left(\tilde{w}_{i,t}^{\xi} \tilde{r}_{i,t}^{1-\xi}\right)^{\gamma} \tilde{P}_{i,t}^{1-\gamma}.
$$

Taking the time difference yields

$$
\hat{x}_{i,t} \equiv \frac{\tilde{x}_{i,t+1}}{\tilde{x}_{i,t}} = \frac{\left(\tilde{w}_{i,t+1}^{\xi} \tilde{r}_{i,t+1}^{1-\xi}\right)^{\gamma} \tilde{P}_{i,t+1}^{1-\gamma}}{\left(\tilde{w}_{i,t}^{\xi} \tilde{r}_{i,t}^{1-\xi}\right)^{\gamma} \tilde{P}_{i,t}^{1-\gamma}} = \left(\hat{w}_{i,t}^{\xi} \hat{r}_{i,t}^{1-\xi}\right)^{\gamma} \hat{P}_{i,t}^{1-\gamma}.
$$

Recall that in the detrended version of the model, the trade flow share from location n to location i at time t is

$$
\lambda_{in,t} = \frac{T^{-\theta} \tilde{A}_{n,t} \left(\kappa_{in,t} \tilde{x}_{n,t}\right)^{-\theta}}{\tilde{P}_{i,t}^{-\theta}},
$$

where T is some constant. Taking the time difference yields

$$
\frac{\lambda_{in,t+1}}{\lambda_{in,t}} = \hat{A}_{n,t+1} \left(\frac{\hat{\kappa}_{in,t+1} \hat{x}_{n,t+1}}{\hat{P}_{i,t+1}}\right)^{-\theta},
$$

which leads to equilibrium condition (E.6).

Note that the detrended price index in location i is

$$
\tilde{P}_{i,t} = T\left(\sum_{n=1}^{N} \tilde{A}_{n,t} \left(\kappa_{in,t} \tilde{x}_{n,t}\right)^{-\theta}\right)^{-1/\theta}.
$$

Taking the time difference, we have

$$\begin{aligned}\hat{P}_{i,t+1} &= \left(\sum_{n=1}^{N} \frac{\tilde{A}_{n,t+1} \left(\kappa_{in,t+1}\tilde{x}_{n,t+1}\right)^{-\theta}}{\sum_{h=1}^{N} \tilde{A}_{h,t} \left(\kappa_{ih,t}\tilde{x}_{h,t}\right)^{-\theta}} \right)^{-1/\theta} \\ &= \left(\sum_{n=1}^{N} \frac{\tilde{A}_{n,t} \left(\kappa_{in}\tilde{x}_{n,t}\right)^{-\theta} \tilde{A}_{n,t+1} \left(\kappa_{in,t+1}\tilde{x}_{n,t+1}\right)^{-\theta} / \tilde{A}_{n,t} \left(\kappa_{in,t}\tilde{x}_{n,t}\right)^{-\theta}}{\sum_{h=1}^{N} \tilde{A}_{h,t} \left(\kappa_{ih,t}\tilde{x}_{h,t}\right)^{-\theta}} \right)^{-1/\theta} \\ &= \left(\sum_{n=1}^{N} \lambda_{in,t} \hat{A}_{n,t+1} \left(\hat{\kappa}_{in,t+1}\hat{x}_{n,t+1}\right)^{-\theta} \right)^{-1/\theta},\end{aligned}$$

where we use $\lambda_{in,t} = \frac{\tilde{A}_{n,t}(\kappa_{in}\tilde{x}_{n,t})^{-\theta}}{\sum_{h=1}^{N}\tilde{A}_{h,t}(\kappa_{ih}\tilde{x}_{h,t})^{-\theta}}$ and which gives equilibrium condition (E.5).

To obtain equilibrium condition (E.7), we use labor market clearing condition (12),

$$\tilde{w}_{i,t+1}L_{i,t+1} = \sum_{n=1}^{N} \lambda_{ni,t+1}\tilde{w}_{n,t+1}L_{n,t+1},$$

and divide by $\tilde{w}_{i,t}L_{i,t}$ on both sides, to obtain

$$\begin{aligned}\hat{w}_{i,t+1}\hat{L}_{i,t+1} &= \frac{1}{\tilde{w}_{i,t}L_{i,t}} \sum_{n=1}^{N} \lambda_{ni,t+1}\tilde{w}_{n,t+1}L_{n,t+1} \\ &= \frac{1}{\tilde{w}_{i,t+1}L_{i,t+1}} \sum_{n=1}^{N} \lambda_{ni,t+1}\hat{w}_{n,t+1}\hat{L}_{n,t+1}\tilde{w}_{n,t}L_{n,t},\end{aligned}$$

where as before we use $\tilde{L}_{n,t} = L_{n,t}$.

Equation (E.8) is exactly the detrended law of motion of capital as in equation (15). To obtain equation (E.9), we use the equilibrium condition:

$$\frac{\tilde{w}_{i,t}\tilde{L}_{i,t}}{\left[\tilde{R}_{i,t} - (1-\delta)\right]\tilde{P}_{i,t}\tilde{K}_{i,t}} = \frac{\xi}{1-\xi}.$$

Taking the time difference and rearranging this expression yields the desired result.

Finally, to obtain the law of motion of knowledge in relative time changes (E.10), note that equation (16) gives the detrended law of motion of knowledge:

$$\tilde{A}_{n,t+1} - \frac{\tilde{A}_{n,t}}{(1+g_A)} = \frac{\alpha_0\Gamma_\rho}{(1+g_A)} \sum_{i=1}^{N} s_{in,t} \left(\tilde{A}_{i,t}\right)^{\rho_l} \sum_{i=1}^{N} \lambda_{ni,t} \left(\frac{\tilde{A}_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m}.$$

Divided by $\tilde{A}_{n,t}$ on both sides, we have

$$\hat{A}_{n,t+1} = \frac{1}{(1+g_A)} + \frac{\alpha_0\Gamma_\rho}{\tilde{A}_{n,t}(1+g_A)} \sum_{i=1}^{N} s_{in,t} \left(\tilde{A}_{i,t}\right)^{\rho_\ell} \left[\sum_{i=1}^{N} \lambda_{ni,t} \left(\frac{\tilde{A}_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m}\right].$$

F Solution Algorithm

In this section of the appendix, we describe the algorithm used to compute the dynamic spatial growth model.

F.1 Algorithm to Solve for the Sequential Equilibrium Given Initial Conditions

In what follows, we describe the algorithm to solve the detrended model given an initial allocation of the economy, $\left(\{L_{i,0}\}_{i=1}^{N}, \{\tilde{K}_{i,0}\}_{i=1}^{N}, \{\lambda_{in,0}\}_{i,n=1}^{N}, \{\mu_{in,-1}\}_{i,n=1}^{N}, \{\tilde{A}_{i,0}\}_{i=1}^{N}\right)$, and given an unanticipated convergent sequence of changes in fundamentals, $\left\{\{\hat{m}_{in,t}\}_{i,n=1}^{N}, \{\hat{\kappa}_{in,t}\}_{i,n=1}^{n}\right\}_{t=1}^{\infty}$. We first describe the algorithm to solve the model under the given initial conditions and constant fundamentals going forward; namely, with $\left\{\{\hat{m}_{in,t}=1\}_{i,n=1}^{N}, \{\hat{\kappa}_{in,t}=1\}_{i,n=1}^{n}\right\}_{t=1}^{\infty}$. We then describe how to solve the model under a change in fundamentals.

1. Initiate the algorithm at $t=0$ with a guess for the path of $\left\{\hat{u}_{i,t+1}^{(0)}\right\}_{t=0}^{T}$, where the superscript (0) indicates that it is a guess. The path should converge to $\hat{u}_{i,T+1}^{(0)}=1$ for sufficiently large T.

2. For all $t \geq 0$, use $\left\{\hat{u}_{i,t+1}^{(0)}\right\}_{t=0}^{T}$ and $\{\mu_{in,-1}\}_{i,n=1}^{N}$ to solve for the path of $\left\{\{\mu_{in,t}\}_{i,n=1}^{N}\right\}_{t=0}^{T}$ using equation (E.2).

3. Use the path for $\left\{\{\mu_{in,t}\}_{i,n=1}^{N}\right\}_{t=0}^{T}$ and $\{L_{i,0}\}_{i=1}^{N}$ to obtain the path for $\left\{\{L_{i,t+1}\}_{i=1}^{N}\right\}_{t=0}^{T}$ using equation (E.3).

4. Solve for the trade equilibrium:

    (a) For each $t \geq 0$, given $\hat{L}_{i,t+1}$, define the term $\hat{\omega}_{i,t}=\tilde{w}_{i,t}^{\xi}\tilde{r}_{i,t}^{1-\xi}$. Guess a value for $\hat{\omega}_{i,t+1}$.

    (b) Obtain $\hat{x}_{i,t+1}$, $\hat{P}_{i,t+1}$, $\lambda_{in,t+1}$, $\tilde{R}_{i,t+1}$, $\tilde{K}_{i,t+1}$, and $\hat{A}_{i,t+1}$ using equations (E.4), (E.5), (E.6), (E.8) and (E.9). Use the fact that $\hat{r}_{i,t+1}=\hat{w}_{i,t+1}\hat{L}_{i,t+1}/\hat{K}_{i,t+1}$ and $\hat{w}_{i,t+1}=\hat{\omega}_{i,t+1}\left(\hat{K}_{i,t+1}/\hat{L}_{i,t+1}\right)^{1-\xi}$.

    (c) Check if the market clearing condition (E.7) holds using $\hat{w}_{i,t+1}=\hat{\omega}_{i,t+1}\left(\hat{K}_{i,t+1}/\hat{L}_{i,t+1}\right)^{1-\xi}$. If it does not, go back to step (a) and adjust the initial guess for $\hat{\omega}_{i,t+1}$ until labor markets clear.

    (d) Repeat steps (a) through (d) for each period t and obtain paths for $\{\hat{w}_{i,t+1}, \hat{P}_{i,t+1}\}_{t=0}^{T}$ for all i.

5. For each t, use $\mu_{in,t}$, $\hat{w}_{i,t+1}$, $\hat{P}_{i,t+1}$, and $\hat{u}_{n,t+2}^{(0)}$ to solve backwards for $\hat{u}_{i,t+1}^{(1)}$ using equation (E.1). This solution delivers a new path for $\left\{\left\{\hat{u}_{i,t+1}^{(1)}\right\}_{i=1}^{N}\right\}_{t=0}^{T}$, where the superscript 1 indicates an updated value for $\hat{u}$.

6. Check whether $\left\{\left\{\hat{u}^{(1)}_{i,t+1}\right\}_{i=1}^{N}\right\}_{t=0}^{T} \approx \left\{\left\{\hat{u}^{(0)}_{i,t+1}\right\}_{i=1}^{N}\right\}_{t=0}^{T}$. If it does not, go back to step 1 and update the initial guess with $\left\{\left\{\hat{u}^{(1)}_{i,t+1}\right\}_{i=1}^{N}\right\}_{t=0}^{T}$.

F.2 Solving for Counterfactual Changes in Fundamentals

We now describe how to solve the dynamic spatial growth model given an unanticipated convergent sequence of changes in fundamentals, $\hat{\Theta}_{t+1} = \left\{\{\hat{m}_{in,t}\}_{i,n=1}^{N}, \{\hat{\kappa}_{in,t}\}_{i,n=1}^{n}\right\}_{t=1}^{\infty}$.

The algorithm used to solve for a change in fundamentals follows the same steps described in the previous section, but the sequence of changes in fundamentals is fed into the model. The main difference from the previous section is that we now consider the fact that agents are surprised in the first period by the changes in fundamentals. The surprise in the changes in fundamentals is captured in the initial gross flow equation. That is, in the first period we now use the following equilibrium condition:

$$\mu_{in,1}(\hat{\Theta}) = \frac{\vartheta_{in,0}\left(\hat{u}_{n,2}(\hat{\Theta})\right)^{\beta/\nu}\left(\hat{m}_{in,1}\right)^{-1/\nu}}{\sum_{i=1}^{N}\vartheta_{ih,0}\left(\hat{u}_{h,2}(\hat{\Theta})\right)^{\beta/\nu}\left(\hat{m}_{ih,1}\right)^{-1/\nu}},$$

where $\vartheta_{in,0} = \mu_{in,0}\exp\left(V_{n,1}(\hat{\Theta}) - V_{n,1}\right)^{\beta/\nu}$. Therefore, we also use the equilibrium condition,

$$\log(\hat{u}_{i,1}) = \log\left(\hat{w}_{i,1}/\hat{P}_{i,1}\right) + \nu\log\left(\sum_{n=1}^{N}\vartheta_{in,0}\left(\hat{u}_{n,2}\right)^{\beta/\nu}\left(\hat{m}_{in,1}\right)^{-1/\nu}\right).$$

G Data Sources, Initial Stock of Knowledge, and Empirical Moments

In this section of the appendix, we describe in more detail the data sources and construction used in the quantitative analysis.

We obtain GDP, employment, export, and import data from the China Compendium of Statistics, 1949-2008.[1] The book consists of three main parts. The first part contains data at the national level compiled by National Bureau of Statistics. The second part presents data from provinces, autonomous regions, and municipalities under the direct jurisdiction of the central government; the data are complied by local statistical bureaus. The third part provides data from the Special Administrative Regions (SARs) of Hong Kong and Macao that have been edited by the National Bureau of Statistics. The national GDP, employment, and trade data do not include those of the Hong Kong SAR, Macao SAR, or Taiwan Province.

[1] The digitized data can be extracted from https://data.cnki.net/area/yearbook/Single/N2010042091?dcode=D03.

G.0.1 List of Provinces

The geographic units used in the quantitative analysis are Chinese provinces and the rest of the world. Strictly speaking, the province-level administrative divisions in China include provinces, autonomous regions, and municipalities under the direct jurisdiction of the central government. For simplicity, we call provinces to these highest-level administrative divisions of China. These provinces are Beijing, Tianjin, Hebei, Shanxi, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangdong, Guangxi, Hainan, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang.

G.0.2 Gross Migration Flows

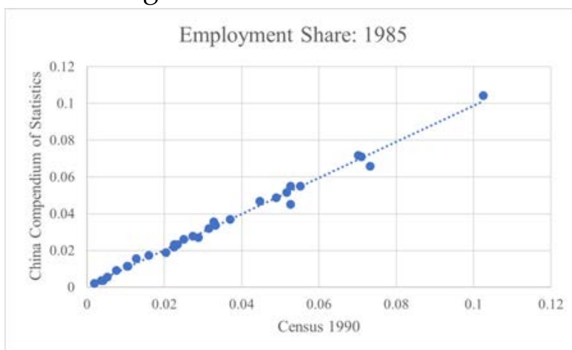
We use the Chinese census data from IPUMS to construct the migration flow matrix. Our constructed migration flow matrix matches the employment share of each province reported in the China Compendium of Statistics. We leverage the 1% samples of the 1990 and 2000 censuses from IPUMS as our data source to calculate migration flows for 1985-1990, 1990-1995, and 1995-2000.

To construct the migration flows for 1985-1990, we proceed as follows. With the 1% sample of the 1990 census, we include the working-age (15-64) population in our sample. Furthermore, we keep any census respondent who is actively employed in 1990. We put a weight on each province to match exactly the provincial employment share shown in the census with that of each province reported by the China Compendium of Statistics. For each individual, we determine the Hukou location as follows. For the 1990 census, the status and nature of registration was asked. If the person chose "(1) residing and registered here", we use the person's location in 1990 as the registration location; if the person chose "(2) residing here over 1 year, but registered elsewhere", "(3) living here less than 1 year absent from registration place over 1 year" or "(4) living here with registration unsettled", we use the person's location in 1985 as the registration location. For a person whose Hukou registration is in category (2)-(4) but who lived in the same province (stayer) in 1985 and 1990, we assign a Hukou place to them as follows. We first construct a sample of migrants who switched their habitant province between 1985 and 1990, as measured in the data. Then, for each destination province, we compute the share of migrants coming from different origin provinces. We assign the Hukou place to the aforementioned stayer according to this share.[2] For each Hukou location (province-level), we construct a five-year migration flow matrix from origin province to destination province. Combining the migration matrix and the data in 1990, we can check whether the employment share of each province out

[2] A potential concern is step migration, i.e., a person does not directly migrate from her registration location to the current location. We cannot check this using the 1990 census. Imbert, Seror, Zhang, and Zylberberg (2022) uses 2005 mini census data to show that step migration was negligible in 2000-2005. We do not expect this to be any different for the period 1985 to 1990.

of the nationwide total employment is consistent with the data from the China Compendium of Statistics (see Figure G.1).

Figure G.1: Data validation



Using the same method, we calculate migration flow between 1995 and 2000 using the 1% sample of the 2000 census from IPUMS. To the best of our knowledge, there are no publicly available micro-level people census data from 1995. When necessary, we thus use as proxy the migration flows from 1985 to 1990 for the flows from 1990 to 1995.

G.0.3 Trade and Production Data

To obtain the trade and production data for China and the rest of the world, we proceed as follows. To take the model to the data, in addition to the data described in the previous section of this appendix, the bilateral trade shares $\lambda_{in,t}$, total expenditure $X_{i,t}$,value added $w_{i,t}L_{i,t} + r_{i,t}K_{i,t}$, and the initial capital stock $K_{i,0}$ must be obtained. We also need to compute the share of value added in gross output, γ, and the share of labor in value added, ξ.

GDP, Employment, Export, and Import Data. We obtain GDP, employment, export, and import data from the China Compendium of Statistics, 1949-2008.

We make several adjustments to the data. First, the Chinese national accounts are based on data provided by local governments to the National Bureau of Statistics (Bai, Hsieh, and Qian (2006), and Chen, Chen, Hsieh, and Song (2019)). Given the incentive of local governments to overstate the local GDP and other measurement discrepancies, the National Bureau of Statistics adjusts the data reported from the local governments to calculate the national-level GDP using independent data sources. Consequently, the reported aggregate GDP is generally lower than the sum of reported province-level GDPs. We address this issue by scaling down province-level GDP by the same proportion for each Chinese province to match the reported GDP at the national level. We follow the same strategy to adjust province-level employment, export, and import data to match their reported national aggregates.

Second, we account for the changing status of Chongqing. Before 1997, Chongqing was not considered a municipality under the direct jurisdiction of the central government. One of the focuses of this paper is to understand the rise of China in 1990s. For most of this period, Chongqing was still part of Sichuan, we thus treat Chongqing and Sichuan as an integrated province, Sichuan-Chongqing, throughout our paper. We aggregate relevant variables for the two regions.

Third, for some provinces, the measurement units are not consistent with the those of the national aggregates. For example, the export and import data of Guangdong Province are inaccurately reported by the local statistical bureau in units of 100 million Chinese Yuan, although the indicated unit is still 10,000 Chinese Yuan. We carefully checked and addressed this type of issues in the data.

We use the GDP deflator from the World Development Indicators compiled by the World Bank, to compute the real GDP at province level at 1990 prices.

To construct the production data for the rest of the world, the PWT 10.0 reports real GDP at constant 2017 national prices (rgdpna) and employment (emp). We rely on the Penn World Table 10.0 (PWT 10.0) to construct data for the rest of the world. We first keep all countries but China. Second, we drop countries with missing data for either GDP or employment. We aggregate all countries in our sample to obtain GDP and employment for the rest of the world. The World Development Indicators database reports the world GDP deflator from 1985 to 2017. Combining the two data sources, we compute GDP for the rest of the world at current year prices and real GDP at 1990 prices.

Capital stock. We follow Shan (2008) to estimate province-level capital stock from 1952 to 2010. We use the perpetual inventory method to estimate the time series of capital stock. For capital stock at the base year, we follow Young (2003), using 10 percent of the gross capital formation in 1952. As Young (2003) and Bai, Hsieh, and Qian (2006) argue, the most appropriate measure of investment in China is fixed capital formation. We obtain this measure from the China Compendium of Statistics. The investment price deflator is constructed by Shan (2008) based on official statistics. We follow Shan (2008) to choose the value for the depreciation rate.

For the rest of the world, we obtain capital stock at constant 2017 national prices from the PWT. We deflate country-level capital stock to reflect 1990 national prices using the GDP deflator. We further adjust the capital stock of the rest of the world by matching the percentage gross fixed capital formation in GDP compiled by the World Bank. We start from the aggregate capital stock of all countries (including China) in 1985 according to the PWT. We adjust for the aggregate capital stock in the years 1990, 1995, and 2000 to match the average gross capital formation (percentage of GDP) in 1985-1990, 1990-1995, and 1995-2000, respectively. Afterward, by excluding the capital stock of China, we obtain the capital stock for the rest of the world.

Shares. We compute the values of $\gamma = 0.38$ and $\xi = 0.54$, which correspond to the parameter values for the year 1990 from world's aggregates in the Eora multi-region input-output table.

To unify the units of measure, GDP, exports, and imports are measured in 100 million USD, while employment is measured in units of 10,000 people.

G.1 Initial Stock of Knowledge

In this section of the appendix, we describe the computation of the initial stock of knowledge across locations. We start from the domestic expenditure $\lambda_{nn,0} = A_{n,0}\left(\frac{x_{n,0}}{P_{n,0}/T}\right)^{-\theta}$. Using this equation, we obtain

$$A_{n,0} = \left(\frac{B\left(w_{n,0}^{\xi} r_{n,0}^{1-\xi}\right)^{\gamma} P_{n,0}^{1-\gamma}}{P_{n,0}/T}\right)^{\theta} \lambda_{nn,0}.$$

Using the first-order condition of the firm's problem, $\frac{w_{n0}L_{n0}}{r_{n0}K_{n0}} = \frac{\xi}{1-\xi}$, we obtain

$$A_{n,0} = (BT)^{\theta}\left(\frac{1-\xi}{\xi}\right)^{(1-\xi)\gamma\theta}\left(\frac{\frac{w_{n,0}L_{n0}}{P_{n,0}}}{(K_{n,0})^{1-\xi}(L_{n,0})^{\xi}}\right)^{\gamma\theta} \lambda_{nn,0}.$$

Finally, using the fact that $w_{n,0}L_{n0} = \xi\left(w_{n,0}L_{n,0} + r_{n,0}K_{n,0}\right)$, we find that the initial stock of knowledge across locations is given by

$$A_{n,0} = \Upsilon\left(\frac{Real\ GDP_{n,0}}{(K_{n,0})^{1-\xi}(L_{n,0})^{\xi}}\right)^{\gamma\theta} \lambda_{nn,0},$$

where $\Upsilon = (BT)^{\theta}(1-\xi)^{(1-\xi)\gamma\theta}(\xi)^{\xi\gamma\theta}$.

G.2 Empirical Moments

In this section of the appendix, Table G.1 presents the empirical moment conditions targeted to discipline the elasticities that govern innovation and idea diffusion $(\alpha_0, \rho_l, \rho_m)$, and the model-implied moments predicted by the evolution of fundamental productivity using equation (9).

Table G.1: Moment Conditions

Moment	Data	Model
Moment 1	164.5	114.5
Moment 2	0.84	0.60
Moment 3	18.5	2.8
Moment 4	5.3	7.1
Moment 5	-8.0	-9.1

Note: The table presents the five moment conditions in the data and the model-implied moment conditions used to discipline the innovation and idea diffusion parameters (α_0, ρ_l, ρ_m). Moment 1 is the average change in fundamental productivity levels across locations. Moment 2 is the average growth rate in fundamental productivities. Moment 3 is the variance in the time changes in fundamental productivity levels (in thousands). Moment 4 is the covariance between the initial fundamental productivities and the change in fundamental productivity levels (in thousands). Moment 5 is the covariance between the initial fundamental productivities and the growth rate in fundamental productivities.

H Additional Quantitative Results

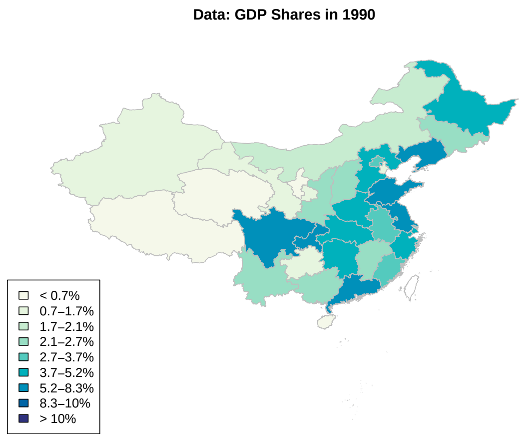
In this section of the appendix, we describe additional results from our quantitative analysis.

H.1 Regional Distribution of Economic Activity

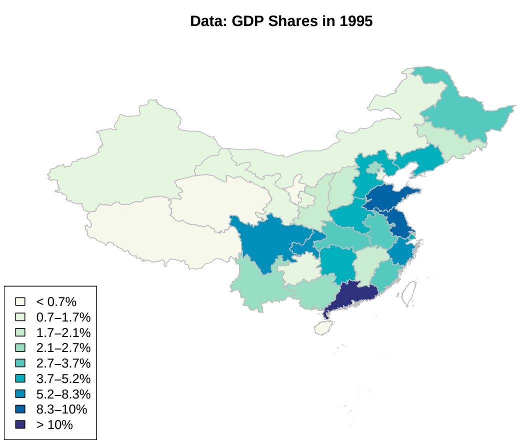
Figures H.1 and H.2 display the evolution of actual GDP shares in China and their evolution under initial 1990 conditions. The figure presents the GDP shares across provinces in China every five years during the period 1990-2010.

Figure H.1: Regional distribution of economic activity (GDP shares)

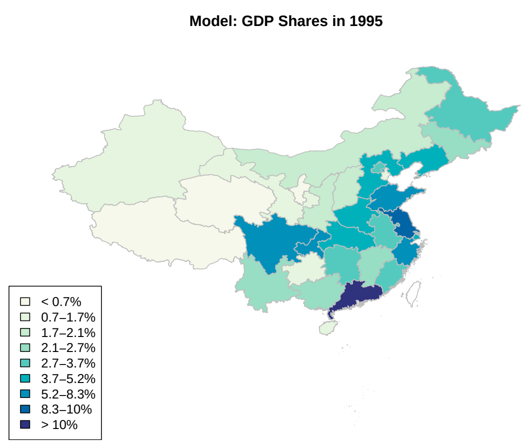
a) 1990



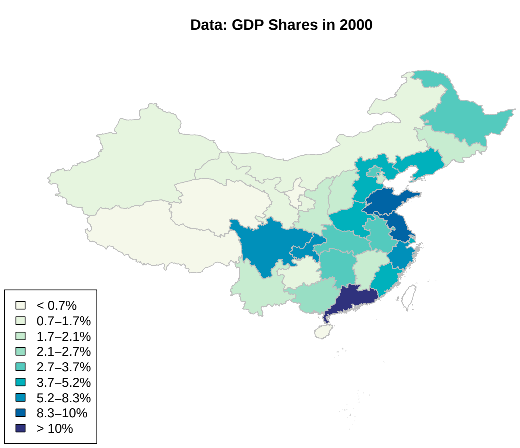
b) Actual
1995



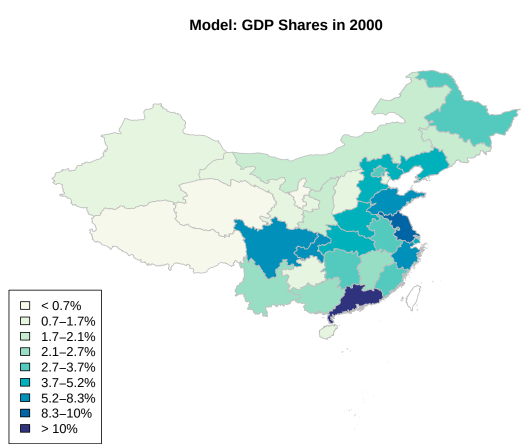
c) 1990 conditions
1995



d) Actual
2000



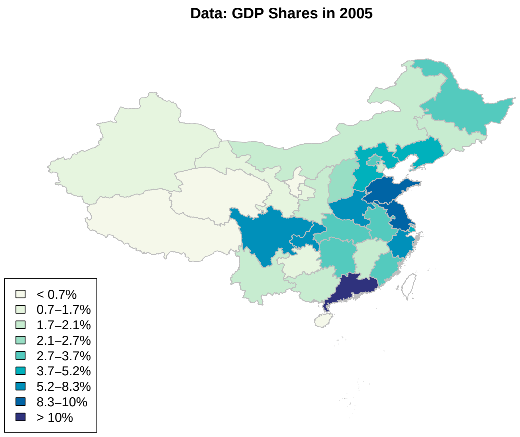
d) 1990 conditions
2000



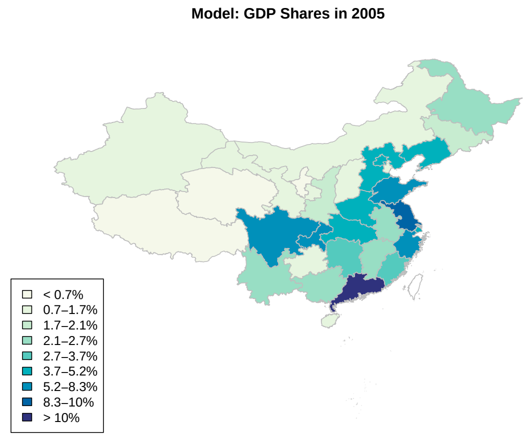
Note: The figures show the distribution of economic activity across provinces in China, measured in GDP shares, in the data and with 1990 fundamentals over the period 1990-2010.

Figure H.2: Regional distribution of economic activity (GDP shares, continued)

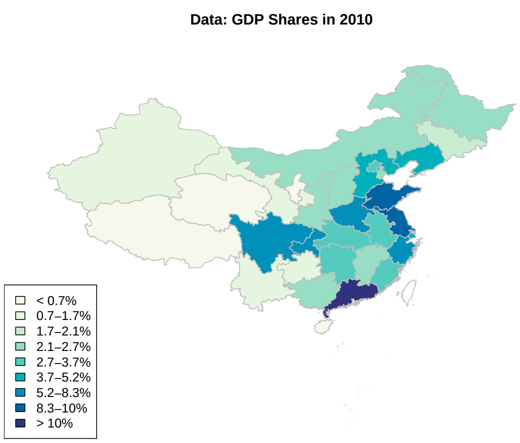
e) Actual
2005



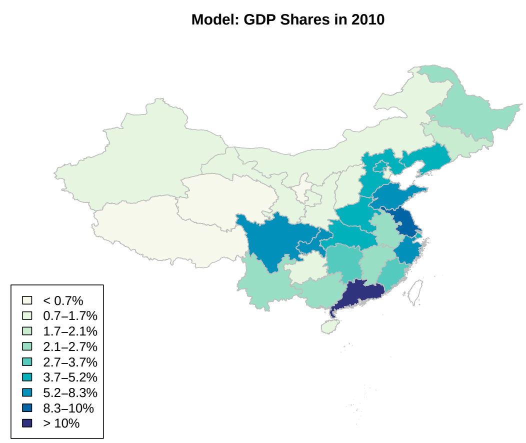
f) 1990 fundamentals
2005



c) Actual
2010



e) 1990 fundamentals
2010

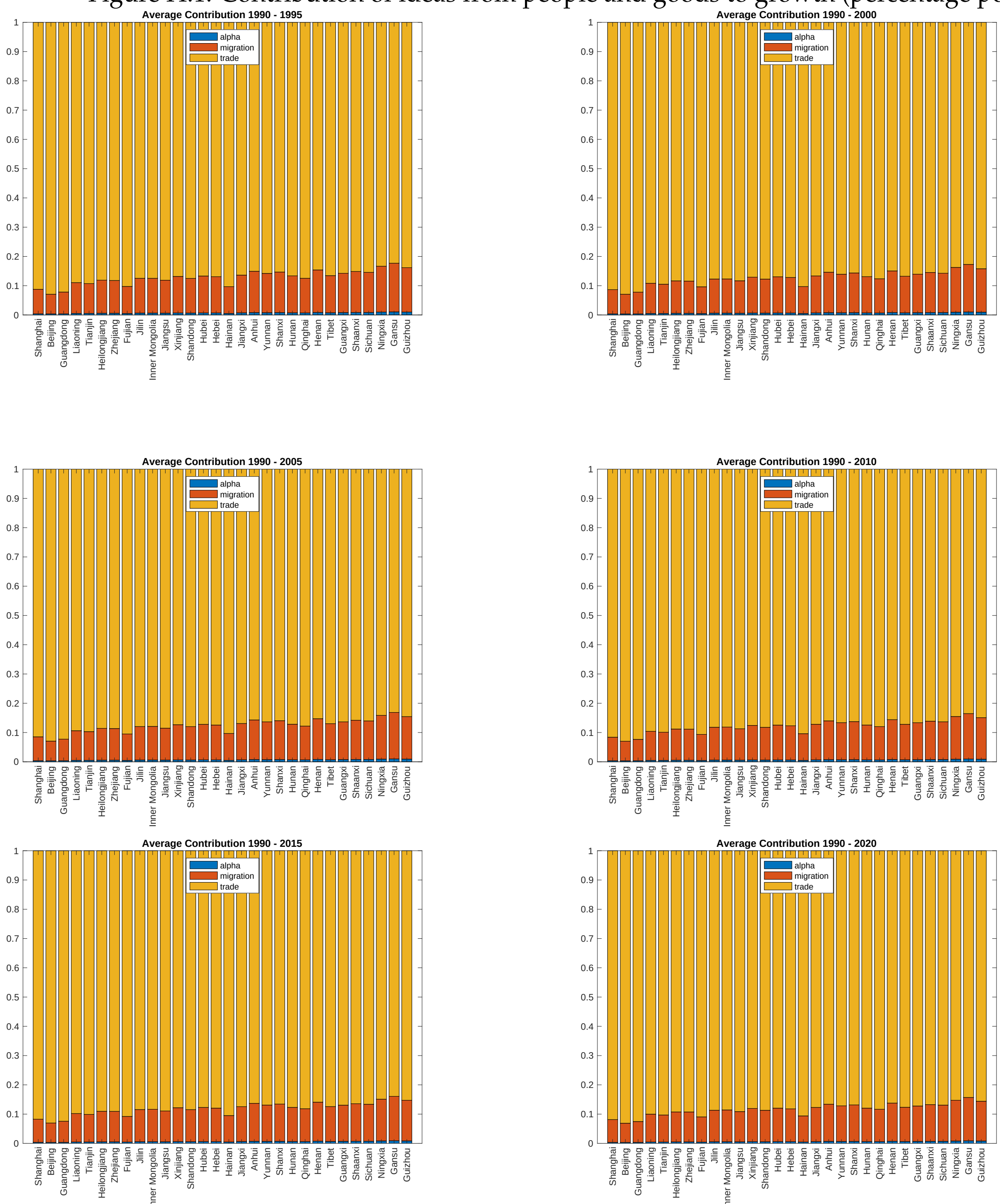


Note: The figures show the distribution of economic activity across provinces in China, measured in GDP shares, in the data and with 1990 fundamentals over the period 1990-2010.

H.2 Spatial Growth from Idea Diffusion

Figure H.1 presents the relative contribution of ideas from people and ideas from goods to the spatial growth across provinces in China over the periods 1990-1995, 1990-2000, 1990-2005, 1990-2010, 1990-2015, and 1990-2020.

Figure H.1: Contribution of ideas from people and goods to growth (percentage points)

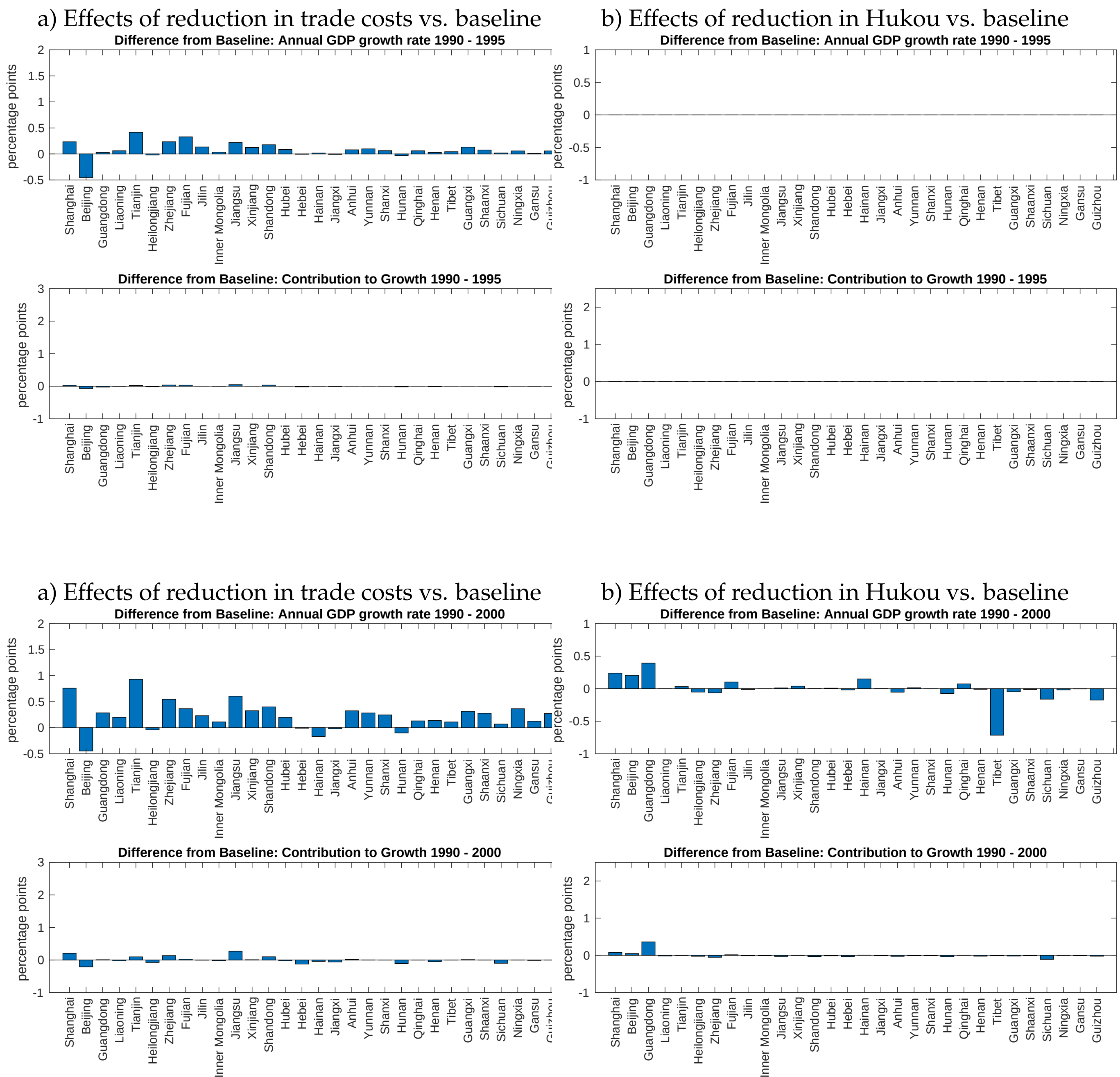


Note: The figures show the relative contribution of ideas from goods, ideas from people, and innovation to growth across provinces with the 1990 initial conditions.

H.3 Spatial Growth Effects of Trade and Hukou Reforms

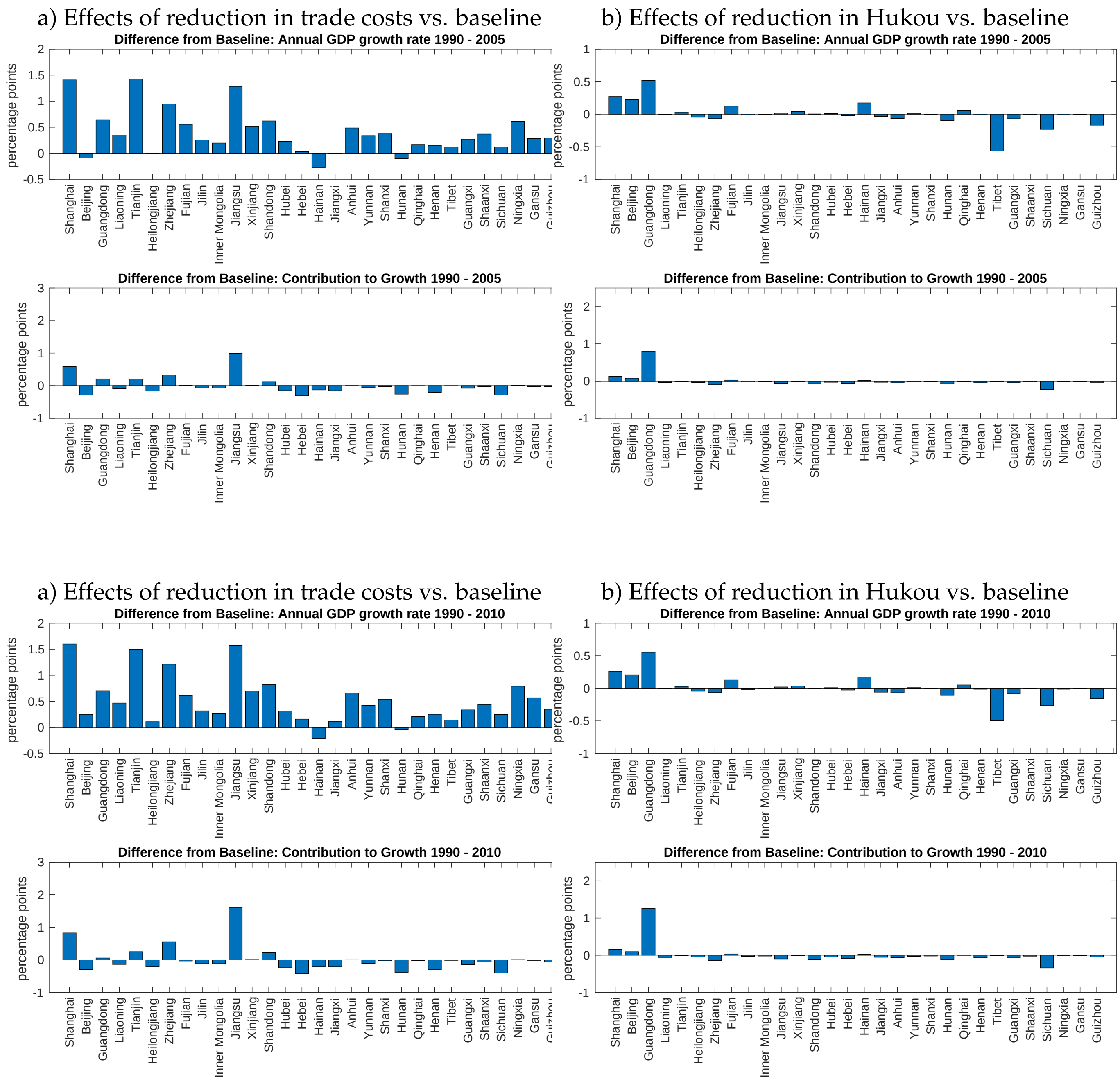
Figures H.2, H.3, and H.4 present the effects of changes in international trade costs and Hukou restrictions relative to the baseline economy with initial conditions in 1990 for alternative time windows.

Figure H.2: Effects of trade and Hukou reforms on spatial growth (percentage points)



Note: The figures show the percentage point change in real GDP growth across provinces as a consequence of the trade and Hukou reforms in different time frames over the period 1990-2020. The left-hand panels present the effects of changes in trade costs, and the right-hand panels show the effects of Hukou. All effects are computed relative to the baseline economy with 1990 trade and migration costs.

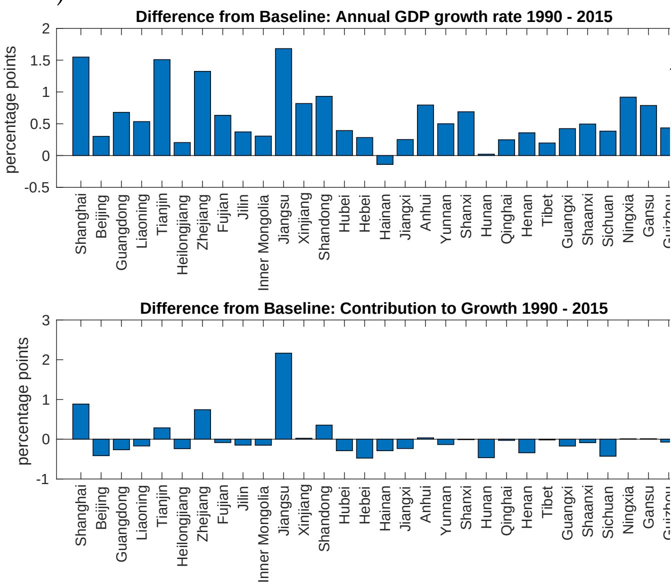
Figure H.3: Effects of trade and Hukou reforms on spatial growth (percentage points)



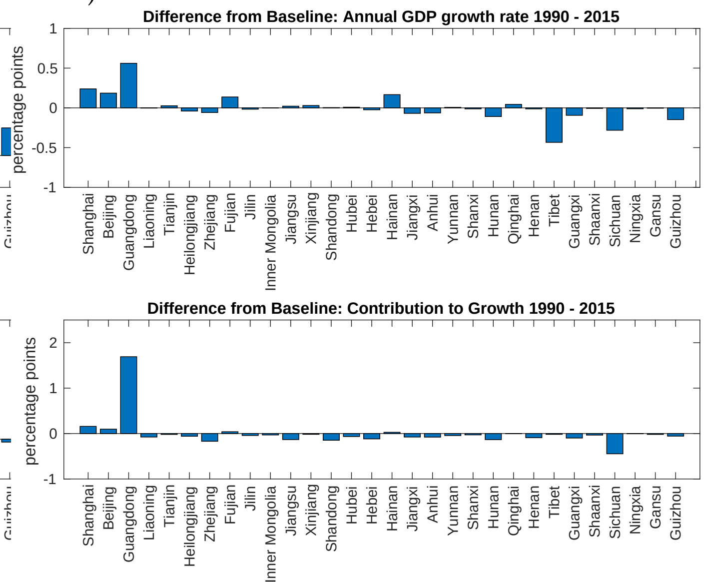
Note: The figures show the percentage point change in real GDP growth across provinces as consequence of the trade and Hukou reforms in different time frames over the period 1990-2020. The left-hand panels present the effects of changes in trade costs, and the right-hand panels show the effects of Hukou. All effects are computed relative to the baseline economy with 1990 trade and migration costs.

Figure H.4: Effects of trade and Hukou reforms on spatial growth (percentage points)

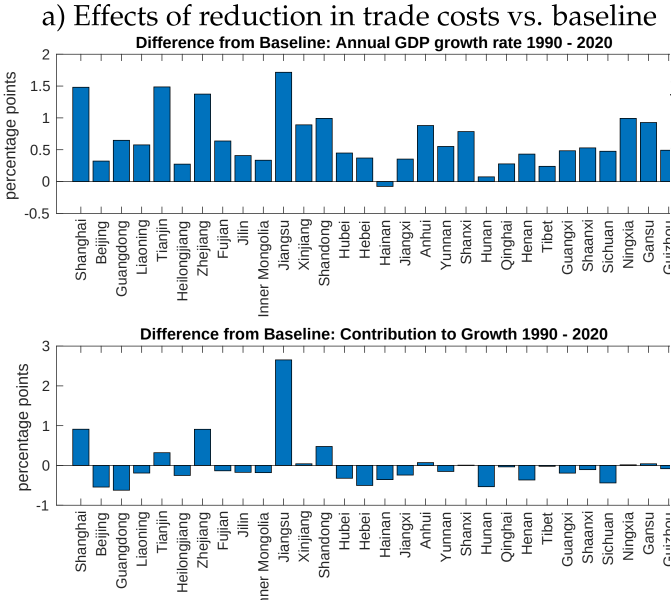
a) Effects of reduction in trade costs vs. baseline



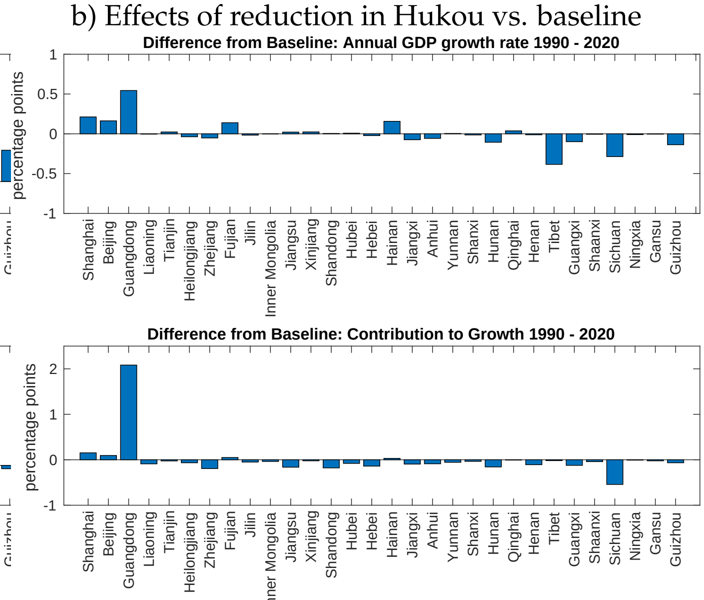
b) Effects of reduction in Hukou vs. baseline



a) Effects of reduction in trade costs vs. baseline



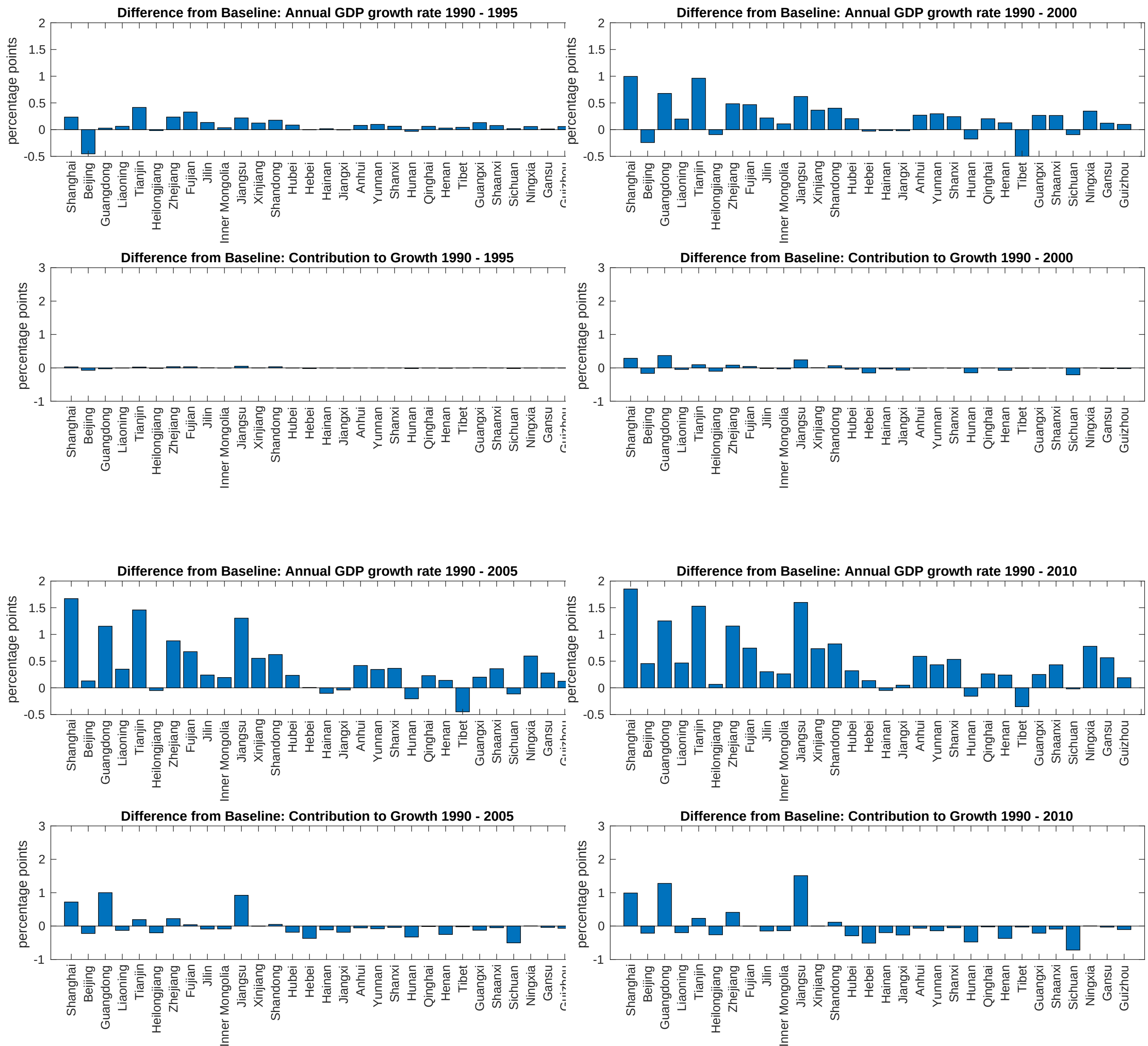
b) Effects of reduction in Hukou vs. baseline



Note: The figures show the percentage point change in real GDP growth across provinces as a consequence of the trade and Hukou reforms in different time frames over the period 1990-2020. The left-hand panels present the effects of changes in trade costs, and the right-hand panels show the effects of Hukou. All effects are computed relative to the baseline economy with 1990 trade and migration costs.

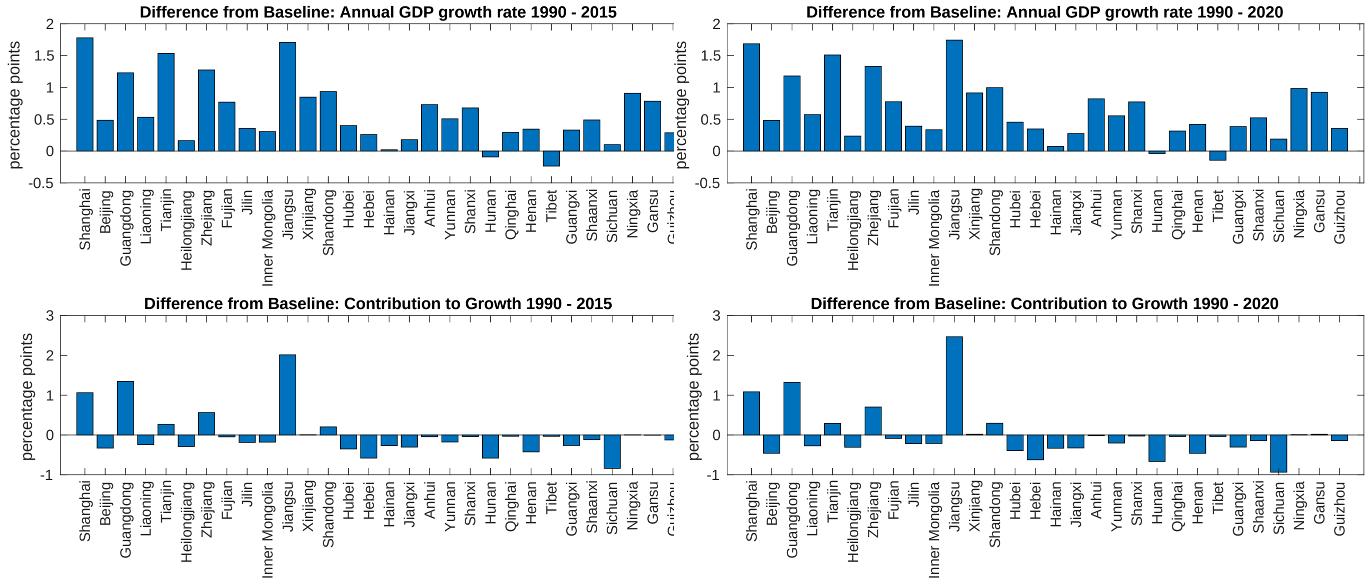
Figures H.5 and H.6 present the combined effects of changes in trade costs and Hukou relative to the baseline economy with 1990 initial conditions. In the upper figures of the panels we display the growth effects, and in the lower figures we present the contribution to aggregate growth relative to the baseline economy. We present results for the time frames 1990-1995, 1990-2000, 1990-2005, 1990-2010, 1990-2015, and 1990-2020.

Figure H.5: Effects of trade and Hukou reforms on spatial growth (percentage points)



Note: The figures show the percentage point change in real GDP growth across provinces as a consequence of the trade and Hukou reforms in different time frames over the period 1990-2020. All effects are computed relative to the baseline economy with 1990 trade and migration costs.

Figure H.6: Effects of trade and Hukou reforms on spatial growth (percentage points)



Note: The figures show the percentage point change in real GDP growth across provinces as a consequence of the trade and Hukou reforms in different time frames over the period 1990-2020. All effects are computed relative to the baseline economy with 1990 trade and migration costs.

I Empirical Evidence of Idea Diffusion

In this section of the appendix, we provide empirical evidence related to the idea diffusion mechanism in our model. In particular, we use province-level patent data, along with trade and migration data, to support the role played by trade and migration in the diffusion of ideas.

We obtain province-level patent data from the China Statistics Yearbooks. There are three types of patents: innovation, utility, and design. For each type of patent, the yearbook reports the number of applications and number of approved patents in a given year. To proxy the measure of knowledge stock, $A_{n,t}$, we calculate the cumulative approved patents of all three types at the province level for each year, starting from 1985. We then compile the province-level knowledge stock in 1985, 1990, 1995, 2000, 2005, and 2010, with which we calculate the change in the knowledge stock every five years from 1985 to 2010. For the approved patents in the rest of the world, we obtain data from Google Patent from 1985-2010 following Liu and Ma (2021).[3]

With the measure for knowledge stock in hand, we construct two diffusion variables. We define idea diffusion through migration as

$$\log\left(migration_{n,t}\right) = \log\left[\sum_{i=1}^{N} s_{in,t} A_{i,t}\right], \tag{I.1}$$

which equals the weighted average of the knowledge stock diffusing to location n at time t via

[3]We are grateful to Song Ma for sharing the Google Patent data.

all migrants and locals. We define diffusion through non-locals (immigration):

$$\log\left(immigration_{n,t}\right) = \log\left[\sum_{i\neq n} s_{in,t} A_{i,t}\right]. \tag{I.2}$$

The way we construct the diffusion through migration shares some similarity with Aghion, Dechezlepretre, Hemous, Martin, and Van Reenen (2016), who investigate how knowledge spillover takes place through inventors. The individual-level population census data for the year 1995 is not publicly available; the migration flows during 1990 - 1995 are thus likewise unavailable. Therefore, we focus on periods after 1995. In Section I.1, we document the correlation between the growth in knowledge and trade openness as well as diffusion through migration. In Section I.2, we empirically test the model-implied law of motion of the knowledge stock (equation (26)) by implementing an instrumental variable strategy. These sections of the appendix complement Section 5 in the main text that provides reduced-form evidence of the contribution of idea diffusion from trade and migration to local knowledge.

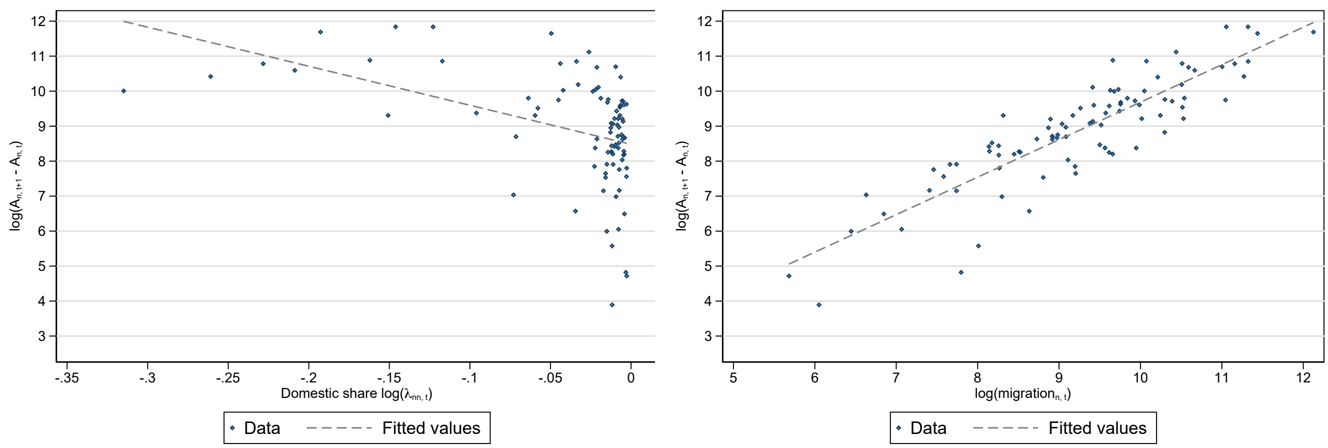
I.1 Simple Correlations

In this section, we compute simple correlations that provide preliminary evidence of knowledge diffusion through trade and migration. In particular, Figure I.1, Panel (a), presents a scatter plot of the change in knowledge stock $\log(A_{n,t+1} - A_{n,t})$, measured as described in the previous section of this appendix, against the domestic expenditure share, $\lambda_{nn,t}$. The negative correlation suggests that the stock of knowledge grows more in locations more open to trade. This correlation is consistent with the mechanism that we highlight in our framework by which global ideas diffuse more to provinces that are more exposed to trade. Figure I.1, Panel (b), shows that the change in knowledge stock is positively correlated with idea diffusion through migration, measured as in equation (I.1), which is also in line with our model of idea diffusion from people.

Figure I.1: Change in the stock of knowledge and idea diffusion

a) Change in the stock of knowledge and trade openness

b) Change in the stock of knowledge and diffusion from migration



Note: The figures show scatter plots of the change in the stock of knowledge against trade openness (Panel (a)) and against diffusion from migration (Panel (b)). The change in the stock of knowledge is measured using patent data, as described in the previous subsection of this appendix; trade openness is measured as the domestic expenditure share, $\lambda_{nn,t}$; and diffusion from migration is measured as in equation (I.1).

I.2 Instrumental Variable Regressions

In this section of the appendix, we use the structure of our model to provide further evidence of our spatial mechanisms using our patent, production, and migration data. Recall that the evolution of the stock of knowledge in our model, according to equation (9), is given by

$$A_{n,t+1} - A_{n,t} = \alpha_t \Gamma_{\rho_\ell,\rho_m} \left[\sum_{i=1}^{N} s_{in,t} \left(A_{i,t}\right)^{\rho_\ell} \right] \left[\sum_{i=1}^{N} \lambda_{ni,t} \left(\frac{A_{i,t}}{\lambda_{ni,t}} \right)^{\rho_m} \right].$$

Taking logs on both sides, we obtain

$$\log\left(A_{n,t+1} - A_{n,t}\right) = \log\left(goods_{n,t}\right) + \log\left(people_{n,t}\right) + \log\alpha_t + \log\Gamma_{\rho_\ell,\rho_m},$$

where $\log\left(goods_{n,t}\right) = \log\left[\sum_{i=1}^{N} \lambda_{ni,t} \left(\frac{A_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m}\right]$ and $\log\left(people_{n,t}\right) = \log\left[\sum_{i=1}^{N} s_{in,t}\left(A_{i,t}\right)^{\rho_\ell}\right]$.

Hence, we specify the following regression equation,

$$\log\left(A_{n,t+1} - A_{n,t}\right) = \beta_m \log\left(goods_{n,t}\right) + \beta_l \log\left(people_{n,t}\right) + \tau_t + \tau + \epsilon_{n,t}, \tag{I.3}$$

where τ_t controls for the time fixed effect, which captures the term $\log\alpha_t$; τ is a constant that captures $\log\Gamma_{\rho_\ell,\rho_m}$; and $\epsilon_{n,t}$ follows an *i.i.d.* standard normal distribution.

It is important to highlight that our model allows for two-way causality in this structural equation due to general equilibrium effects. That being said, in what follows, we still try to establish causality between the growth in knowledge stock and the diffusion through goods and people in order to provide further evidence of our idea diffusion mechanisms. Hence, we try to address potential endogeneity issues in equation (I.3).

As described before, locations more exposed to international trade benefit more from the global diffusion of ideas and experience faster growth in the stock of knowledge. At the same time, fast-growing locations may build up their comparative advantages, impacting international trade as a result. To address potential endogeneity issues, we need an instrument for $\log(goods_{n,t})$. We instrument $\lambda_{ni,t}$ by $\lambda_{ni,1985}$, as the growth prospect can hardly affect the trade pattern 15-25 years before.

For the variable $\log(people_{n,t})$, the reverse causality concern also holds. Locations where the stock of knowledge grows faster might experience more immigration and less outmigration, which affects the knowledge diffusion through people in those locations. Furthermore, a higher share of immigration might lead to faster or slower growth in the stock of knowledge depending on the relative knowledge level and insights of the locals and the immigrants. Hence, the endogeneity issues might lead to either upward or downward estimation bias. To address this, we need an instrumental variable for $\log(people_{n,t})$. We instrument $s_{in,t}$ by $s_{in,1985}$. The rationale is similar to the case of trade; the growth of the stock of knowledge stock can hardly be anticipated by people 15-25 years before, so $s_{in,1985}$ is exogenous from the perspective of location location n in year $t \geq 2000$. In short, the instrument for $\log(goods_{n,t})$ is defined as

$$\log\left[\sum_{i=1}^{N} \lambda_{ni,1985} \left(\frac{A_{i,t}}{\lambda_{ni,1985}}\right)^{\rho_m}\right],$$

and the instrument for $\log(people_{n,t})$ is defined as

$$\log\left[\sum_{i=1}^{N} s_{in,1985} \left(A_{i,t}\right)^{\rho_\ell}\right],$$

where $\rho_m = 0.61$ and $\rho_l = 0.2$ are taken from our GMM estimation. The IV regression results are reported in Table I.2. In Column (1), we directly test our model-implied law of motion of the change in knowledge stock, and we do not control for endogeneity. The positive and significant coefficients of the two diffusion variables are consistent with our spatial mechanisms and in line with the reduced-form evidence presented in Section 5. As the patent stock is a proxy for the knowledge stock in the model and is likely a function of knowledge stock and other factors, the magnitudes of the coefficients do not need to be constrained to be close to one. In Columns (2), we report the instrumental variable regression results.[4] The positive effects of idea diffusion through international trade and migration on the growth in knowledge stock are still salient.

[4] In 1985 Hainan had not been elevated to the status of a province; therefore, Hainan is dropped from the sample in the instrumental variable regressions.

Table I.1: Estimates of the effects of diffusion through trade and migration on knowledge growth

Dept. Var: $\log\left(A_{n,t+1} - A_{n,t}\right)$	OLS	IV
	(1)	(2)
$\log\left(goods_{n,t}\right)$	0.470***	0.390**
	(0.127)	(0.162)
$\log\left(people_{n,t}\right)$	4.996***	5.055***
	(0.238)	(0.300)
Constant	-3.317***	-4.157***
	(0.835)	(1.034)
Kleibergen-Paap rk Wald F statistic		87.54
Observations	90	87
R-squared	0.935	0.935
Year FE	✓	✓

Note: Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1.