Online Supplement for "Unpacking Aggregate Welfare in a Spatial Economy"

Eric Donald, Masao Fukui, Yuhei Miyauchi

D Details on Extensions

D.1 Non-Separable Utility

In the baseline model, we have focused on a specification where preference shocks are additively separable. This section relaxes this assumption.

We now assume that utility in location i is given by $U_i^\theta(C_i^\theta, \epsilon_i^\theta)$. Compared to the additively separable specification, marginal utility in each location now depends on preference shock draws. To see this, the average marginal utility for households deciding to live in location j is given by

$$\mathbb{E}\left[\frac{\partial}{\partial C_j^\theta} U_j^\theta(C_j^\theta, \epsilon_j^\theta) | j = \arg\max_i U_i^\theta(C_i^\theta, \epsilon_i^\theta)\right]. \tag{D.1}$$

Unlike the additively separable specification, where $\frac{\partial}{\partial C_j^\theta} U_j^\theta(C_j^\theta, \epsilon_j^\theta) = u_j^{\theta\prime}(C_j^\theta)$, the selection of preference shocks influences the marginal utility of consumption in each location.

Under this general preference specification, the isomorphic representation of households' location decisions in Lemma 1 is modified as

$$\max_{\{\mu_j^\theta\}_j : \sum_j \mu_j^\theta = 1} \mathcal{U}^\theta(\{\mu_j^\theta\}), \tag{D.2}$$

where

$$\begin{aligned} \mathcal{U}^\theta(\{\mu_j^\theta\}) = & \max_{\{\mathbb{I}_j^\theta(\omega)\}_{\omega,j}} \int_0^1 \sum_j u_j^\theta(C_j^\theta, \epsilon_j^\theta(\omega)) \mathbb{I}_j^\theta(\omega) d\omega \\ \text{s.t.} \quad & \int_0^1 \mathbb{I}_j^\theta(\omega) d\omega = \mu_j^\theta \\ & \sum_j \mathbb{I}_j^\theta(\omega) = 1. \end{aligned} \tag{D.3}$$

We followed the same notation and setup as in Appendix A.1.

Under the additively separable specification, $\mathcal{U}^\theta(\{\mu_j^\theta\}) = \sum_j \mu_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\mu_j^\theta\})$

and $\partial\mathcal{U}^\theta/\partial C_j^\theta = \mu_j^\theta u_j^{\theta\prime}(C_j^\theta)$, so expected marginal utility only depends on j's population and consumption. In the general case, it is affected by the entire population distribution $\{\mu_j^\theta\}_j$ and consumption vector $\{C_j^\theta\}_j$ beyond location j through the selection of preference draws.

In this generalized environment, Proposition 1 is simply modified by replacing the prior marginal utility $u_j^{\theta\prime}(C_j^\theta)$ with the one under this general specification. In particular, (ii) MU dispersion term becomes

$$\Omega_{MU} = \mathbb{E}_\theta\left[\mathrm{Cov}_{j|\theta}\left(-\frac{P_j^\theta}{\mathcal{MU}_j^\theta}, \mathcal{MU}_j^\theta dC_j^\theta\right)\right], \tag{D.4}$$

where

$$\mathcal{MU}_j^\theta = \frac{1}{\mu_j^\theta}\frac{\partial\mathcal{U}^\theta}{\partial C_j^\theta}. \tag{D.5}$$

Conditional on the price normalization (18) using this marginal utility, all other terms are unaffected.

D.2 Idiosyncratic Productivity Shocks

We generalize our baseline model by allowing households to draw idiosyncratic productivity $\boldsymbol{z}^\theta = (z_1^\theta, z_2^\theta, \ldots, z_N^\theta)$, in addition to preference shocks $\boldsymbol{\epsilon}^\theta$. When a household decides to live in location j, the efficiency units of labor that the household supplies is z_j^θ.

We make several modifications to our baseline model to make our analysis tractable and transparent. First, we restrict our attention to the case of log utility,

$$u_j^\theta(c) = B_j^\theta \ln c + D_j^\theta, \tag{D.6}$$

where B_j^θ and D_j^θ are slope and intercept parameters specific to location and household type. Second, we assume that location-specific transfers are linear in household labor income, which we denote as τ_j^θ. Third, we assume away the presence of fixed factors.

The household's location choice problem with productivity draw $\boldsymbol{z}^\theta$ and preference draw $\boldsymbol{\epsilon}^\theta$ is

$$\max_j u_j^\theta(c_j^\theta) + \epsilon_j^\theta \tag{D.7}$$

$$\text{s.t.} \quad P_j^\theta c_j^\theta = z_j^\theta w_j^\theta(1+\tau_j^\theta). \tag{D.8}$$

Let

$$C_j^\theta \equiv \frac{w_j^\theta(1+\tau_j^\theta)}{P_j^\theta} \tag{D.9}$$

denote the consumption of household type θ in location j per efficiency units of labor. With our assumption on the utility function (D.6), we can write the location choice problem as

$$\max_j u_j^\theta(C_j^\theta) + \underbrace{B_j^\theta \ln z_j^\theta + \epsilon_j^\theta}_{\equiv \varepsilon_j^\theta}. \tag{D.10}$$

Viewing ε_j^θ as the convoluted idiosyncratic productivity and amenity shocks, we can apply Lemma 1 to obtain the same location choice characterization as in the baseline model:

$$\max_{\{\mu_j^\theta\}_j : \sum_j \mu_j^\theta = 1} \sum_j \mu_j^\theta u_j^\theta(C_j^\theta) + \psi^\theta(\{\mu_j^\theta\}). \tag{D.11}$$

Let $\hat{\mu}_j^\theta(\{C_i^\theta\})$ be the location choice function associated with the solution to the above problem.

It will be convenient to define the average efficiency units of labor in each location-type pair as a function of a vector of $\{C_i^\theta\}$:

$$Z_j^\theta(\{C_i^\theta\}) \equiv \mathbb{E}\left[z_j^\theta \middle| j = \arg\max_i u_i^\theta(C_i^\theta) + B_i^\theta \ln z_i^\theta + \epsilon_i^\theta \right]. \tag{D.12}$$

To the extent that the location choice function is invertible,[1] i.e., the inverse function of $\hat{\mu}_j^\theta(\{C_i^\theta\})$, $C_j^\theta = \hat{C}_j^\theta(\{\mu_i^\theta\})$ exists, we can alternatively define the average efficiency units of labor as a function of location choice probabilities:

$$\mathcal{Z}_j^\theta(\{\mu_i^\theta\}) = Z_j^\theta(\{\hat{C}_l^\theta(\{\mu_i^\theta\})\}). \tag{D.13}$$

The goods market clearing conditions are modified as follows

$$\sum_\theta c_{ij,k}^\theta + \sum_{l,m} x_{jl,m}^{i,k} = \mathcal{A}_{ij,k} f_{ij,k}(\mathbf{l}_{ij,k}, \mathbf{x}_{ij,k}) \tag{D.14}$$

$$\mathcal{Z}_j^\theta C_j^\theta \ell^\theta \mu_j^\theta = \mathcal{C}_j^\theta(\boldsymbol{c}_j^\theta), \tag{D.15}$$

[1] See Berry, Gandhi, and Haile (2013) for a sufficient condition for invertibility.

where the first equation is modified due to the absence of a fixed factor, and the second equation takes into account heterogeneity in consumption within a location-type pair. The labor market clearing condition is

$$\sum_{i,k} l^{\theta}_{ji,k} = \mathcal{Z}^{\theta}_j \ell^{\theta} \mu^{\theta}_j, \tag{D.16}$$

which takes into account heterogeneity in efficiency units of labor within a location-type pair. The rest of the equilibrium conditions remain unchanged.

It is be straightforward to extend Lemma 2 to this environment. Any decentralized equilibrium solves the following pseudo-planning problem:

$$W = \max_{\{W^{\theta},\{C^{\theta}_j,\mathbf{c}^{\theta}_j,\mu^{\theta}_j\}\},\{\mathbf{x}_{ij,k},\mathbf{l}_{ij,k},h_{ij,k},\mathcal{A}_{ij,k}\}} \mathcal{W}(\{W^{\theta}\}) \tag{D.17}$$

subject to (7), (D.14)-(D.16)

$$W^{\theta} = \sum_j \mu^{\theta}_j u^{\theta}_j(C^{\theta}_j) - \psi^{\theta}(\{\mu^{\theta}_j\}) \tag{D.18}$$

$$\{\mu^{\theta}_j\}_j \in \underset{\{\tilde{\mu}_j\}_j:\sum_j \tilde{\mu}_j=1}{\arg\max} \sum_j \tilde{\mu}^{\theta}_j u^{\theta}_j(C^{\theta}_j) - \psi^{\theta}(\{\tilde{\mu}^{\theta}_j\}) \tag{D.19}$$

$$C^{\theta}_j = \check{C}^{\theta}_j \tag{D.20}$$

Applying the envelope theorem, we obtain

$$\begin{aligned} dW &= \Omega_T + \Omega_{MU} + \Omega_{TE} + \Omega_R \\ &\quad + \sum_{\theta}\sum_j \left[w^{\theta}_j - P^{\theta}_j C^{\theta}_j\right] \mathcal{Z}^{\theta}_j \ell^{\theta} d\mu^{\theta}_j + \sum_{\theta}\sum_j\sum_l \left[w^{\theta}_l - P^{\theta}_l C^{\theta}_l\right] \ell^{\theta}\mu^{\theta}_l \frac{\partial \mathcal{Z}^{\theta}_l}{\partial \mu^{\theta}_j} d\mu^{\theta}_j \end{aligned}$$

Denoting $T^{\theta}_j = \tau^{\theta}_j w^{\theta}_j \mathcal{Z}^{\theta}_j$ as the average transfers that households of type θ in location j receive, we can rewrite the above expression as follows:

$$\begin{aligned} dW &= \Omega_T + \Omega_{MU} + \Omega_{TE} + \Omega_R \\ &\quad + \underbrace{\mathbb{E}_{\theta}\left[\mathrm{Cov}_{j|\theta}(-T^{\theta}_j, d\ln l^{\theta}_j)\right] + \mathbb{E}_{\theta}\left[\mathrm{Cov}_{j|\theta}\left(-\sum_l T^{\theta}_l l^{\theta}_l \frac{\partial \ln \mathcal{Z}^{\theta}_l}{\partial \ln \mu^{\theta}_j}, d\ln l^{\theta}_j\right)\right]}_{\text{(iii) Fiscal Externality } (\Omega_{FE})}. \end{aligned}$$

Therefore the only difference from Proposition 1 is the second term inside the (iii) fiscal externality term. This term arises because migration changes the composition of workers in all locations, which in turn affects the government's budget. For example, suppose that migration into location j is associated with an increase in the average productivity of workers living in location j but a decrease in other locations. If location j is a net taxpayer ($\tau_j^\theta < 0$ and thereby $T_j^\theta < 0$), then this will tend to slacken the government's budget.

D.3 Shocks to Amenity and Amenity Externalities

In Section 3, we analyzed the effects of productivity shocks on aggregate welfare. In some contexts, researchers are interested in shocks to amenities instead of productivity. The analysis in Section 3 allows for this possibility by interpreting some intermediate goods as local amenities. From a measurement perspective, applying Proposition 1 requires knowledge of prices associated with amenities, which are often unobserved. Below, we provide an alternative expression for Proposition 1 without using prices for amenities.

To consider this extension, we explicitly introduce amenities as an argument in the utility function as follows:

$$U_j^\theta(C_j^\theta, \mathcal{B}_j^\theta, \epsilon_j^\theta) = u_j^\theta(C_j^\theta, \mathcal{B}_j^\theta) + \epsilon_j^\theta, \tag{D.21}$$

where $\mathcal{B}_j^\theta$ is the amenity in region j. Furthermore, we assume that amenities take the following form:

$$\mathcal{B}_i^\theta = B_i^\theta g_i^{B,\theta}(\{l_i^{\tilde{\theta}}\}), \qquad \gamma_i^{B,\tilde{\theta}\theta} = \frac{\partial \ln g_i^{B,\theta}}{\partial \ln l_i^{\tilde{\theta}}}, \tag{D.22}$$

where B_i^θ is the fundamental component of amenities, $g_i^{B,\theta}(\{l_i^{\tilde{\theta}}\})$ is the spillover function, and $\gamma_i^{B,\tilde{\theta}\theta}$ is the amenity spillover elasticity from type $\tilde{\theta}$ to type θ in location i.

Under this extension, Proposition 1 is modified as follows. Consider an arbitrary set of small shocks to the exogenous components of productivity $\{d \ln A_{ij,k}\}$ and amenities

$\{d\ln B_i^\theta\}$. The first-order impact on aggregate welfare can be expressed as

$$dW = \underbrace{\sum_{i,j,k} p_{ij,k}y_{ij,k}d\ln A_{ij,k} + \sum_{i,\theta} l_i^\theta \partial_B u_i^\theta \mathcal{B}_i^\theta d\ln B_i^\theta}_{\text{(i) Technology } (\Omega_T)} + \underbrace{\mathbb{E}_\theta\left[\text{Cov}_{j|\theta}\left(-\frac{P_j^\theta}{\partial_C u_j^\theta}, \partial_C u_j^\theta dC_j^\theta\right)\right]}_{\text{(ii) MU Dispersion } (\Omega_{MU})}$$
$$+ \underbrace{\mathbb{E}_\theta\left[\text{Cov}_{j|\theta}(-T_j^\theta, d\ln l_j^\theta)\right]}_{\text{(iii) Fiscal Externality } (\Omega_{FE})} + \underbrace{\mathbb{E}_\theta\left[\text{Cov}_{j|\theta}(\sum_{l,k} p_{jl,k}y_{jl,k}\frac{1}{l_j^\theta}\gamma_{jl,k}^\theta + \sum_{\tilde\theta}\partial_b u_j^{\tilde\theta}\mathcal{B}_j^{\tilde\theta}\gamma_j^{B,\theta\tilde\theta}, d\ln l_j^\theta)\right]}_{\text{(iv) Technological Externality } (\Omega_{TE})}$$
$$+ \underbrace{\text{Cov}_\theta\left(\Lambda^\theta - \mathbb{E}_{j|\theta}\left[\frac{P_j^\theta}{\partial_C u_j^\theta(C_j^\theta)}\right], \mathbb{E}_{j|\theta}\left[\partial_C u_j^\theta dC_j^\theta\right]\right)}_{\text{(v) Redistribution } (\Omega_R)}. \tag{D.23}$$

where $\partial_B u_j^\theta \equiv \frac{\partial u_j^\theta}{\partial B_j^\theta}$ and $\partial_C u_j^\theta \equiv \frac{\partial u_j^\theta}{\partial C_j^\theta}$. The main difference from Proposition 1 is the additional components in the (i) technology and (iv) technological externality terms. The second component inside the (i) technology term captures the effects of exogenous amenity shocks, absent reallocation effects. The coefficient in front of $d\ln B_i^\theta$, $l_i^\theta\partial_B u_i^\theta\mathcal{B}_i^\theta$, is the population-weighted sum of the marginal utility of amenities. This term strongly resembles the technology effect from productivity (the first term). In particular, if amenities are traded and priced in the market, $\partial_B u_i^\theta$ corresponds to the competitive price of the amenity, and hence $l_i^\theta\partial_B u_i^\theta\mathcal{B}_i^\theta$ is the total sales of type θ amenities in location i, corresponding to $p_{ij,k}y_{ij,k}$. The second component inside the (iv) technological externality term has the same feature: if the amenity is traded, the term reflecting changes in amenities from externalities collapses to the same form as the productivity externality term.

D.4 Isomorphism between Amenity Externalities and Preference Shocks

In the quantitative spatial equilibrium literature, researchers often argue that amenity congestion externalities are isomorphic to preference shocks and use these specifications interchangeably.[2] This section discusses this isomorphism through the lens of our framework.

[2]For example, see Allen and Arkolakis (2014) and Desmet, Nagy, and Rossi-Hansberg (2018) for papers that mention the isomorphism between the two specifications.

For expositional convenience, we assume a single type and drop superscript θ. This implies $l_j = \mu_j$. Consider the following utility specification with amenity externalities, rather than preference shocks:

$$U_j(C_j, \mathcal{B}_j) = u_j(C_j) + \mathcal{B}_j, \qquad \mathcal{B}_j = -\frac{1}{\nu} \ln S_j(\{l_i\}), \tag{D.24}$$

where $S_j(\{l_i\})$ satisfies the following property

$$\frac{1}{\nu} \sum_j l_j \frac{\partial \ln S_j}{\partial l_i} = 1. \tag{D.25}$$

Note that this specification accommodates the possibility that the population in location i generates externalities in other regions. A special case of this specification is when $\ln S_j = l_j^\nu$, i.e., congestion is iso-elastic to local population size with elasticity ν.

In an interior equilibrium, utility levels are equalized across all locations:

$$u_j(C_j) - \frac{1}{\nu} \ln S_j = \bar{u}, \tag{D.26}$$

for some $\bar{u}$.

Now we show that this is isomorphic to the case where there are no amenity externalities but preference shocks follow a max-stable multivariate Gumbel distribution with shape parameter ν. That is,

$$U_j(C_j, \epsilon_j) = u_j(C_j) + \epsilon_j, \tag{D.27}$$

where $\{\epsilon_j\}$ follows specification (39). As we show in Appendix C, with a multivariate Gumbel distribution, $\psi(\{l_j\})$ in Lemma 1 takes the form $\psi(\{l_j\}) = \frac{1}{\nu} \sum_j l_j \ln S_j(\{l_i\})$, where $S_j(\{l_i\})$ satisfies (D.25). The first-order condition for the representative household problem is

$$u_j(C_j) - \frac{1}{\nu} \ln S_j - \underbrace{\frac{1}{\nu} \sum_i l_i \frac{\partial \ln S_i}{\partial l_j}}_{=1} = \delta, \tag{D.28}$$

which is the same as (D.26). Therefore, the equilibrium allocations will be identical. Moreover, it is also straightforward to see that both specifications deliver the same expected utility, thereby delivering identical normative predictions as well.

This isomorphism arises because this particular form of congestion externality does not induce misallocation. In particular, the amenity component of the (iv) technological externality term in Equation (D.23) comes down to

$$\mathrm{Cov}_j\left(-\sum_i l_i \frac{\partial \ln S_i}{\partial \ln l_j}, d\ln l_j\right) = \mathrm{Cov}_j(-\nu, d\ln l_j) = 0, \tag{D.29}$$

where we used (D.25) and $\partial_b u_i = 1$. Given that all other terms in Equation (D.23) are identical between the two specifications, the aggregate welfare predictions are also isomorphic.

This discussion also clarifies that this isomorphism only holds when preference shocks follow a max-stable multivariate Gumbel distribution, or equivalently, when the congestion externality takes the specific functional form given by (D.24) and (D.25). Outside of these special cases, congestion externalities generate misallocation and the isomorphism does not hold in general.[3]

D.5 Non-Welfarist Welfare Criteria

Consider a general non-welfarist welfare objective

$$W = \mathcal{W}(\{\mathcal{U}^{SP,\theta}(\{C_j^\theta, \mu_j^\theta\})\}),$$

where $\mathcal{U}^{SP,\theta}$ is defined arbitrarily on the population distribution and consumption of type θ households. Then, by applying the envelope theorem to the pseudo-planning problem as in the proof of Proposition 1, we have

$$dW = \sum_{i,j,k} p_{ij,k} y_{ij,k} d\ln A_{ij,k} + \sum_\theta \sum_j \left[\ell^\theta \Lambda^\theta \frac{\partial \mathcal{U}^{\mathcal{SP},\theta}}{\partial C_j^\theta} - l_j^\theta P_j^\theta\right] dC_j^\theta$$

[3] Fajgelbaum and Gaubert (2020) show that under a multiplicative utility specification, spatial equilibria involve misallocation even with iso-elastic amenity externalities. Through the lens of Equation (D.23), this source of misallocation appears in the (ii) MU dispersion term. Multiplicative amenities without preference shocks imply that the marginal utility of income is not equalized across locations. Furthermore, unlike in our baseline model which abstracts from direct effects of shocks on utility, the utility changes from consumption changes dC_j are not equalized because of the changes in utility from amenities. Therefore, the (ii) MU dispersion term is not zero. Note that this specification is isomorphic to the specification with multiplicative max-stable Fréchet shocks (without amenity externality) as discussed in Section 4.1. In that case, the dispersion of marginal utility instead arises from preference shock draws.

$$+\sum_{\theta}\sum_{j}\ell^{\theta}\Lambda^{\theta}\frac{\partial\mathcal{U}^{\mathcal{SP},\theta}}{\partial\mu_j^{\theta}}d\mu_j^{\theta}+\sum_{\theta}\sum_{j}\left[w_j^{\theta}-P_j^{\theta}C_j^{\theta}\right]dl_j^{\theta}$$

$$+\sum_{\theta}\sum_{j}\sum_{l,k}p_{jl,k}y_{jl,k}\frac{1}{l_j^{\theta}}\gamma_{jl,k}^{\theta}dl_j^{\theta} \tag{D.30}$$

The differences from our main proposition are in the second and the third terms, which we can rewrite as

$$\sum_{\theta}\ell^{\theta}\sum_{j}\left[\Lambda^{\theta}\frac{\partial\mathcal{U}^{\mathcal{SP},\theta}}{\partial C_j^{\theta}}-P_j^{\theta}\mu_j^{\theta}\right]dC_j^{\theta}$$

$$=\sum_{\theta}\ell^{\theta}\sum_{j}\mu_j^{\theta}\left[\Lambda^{\theta}\left(\frac{\frac{1}{\mu_j^{\theta}}\frac{\partial\mathcal{U}^{\mathcal{SP},\theta}}{\partial C_j^{\theta}}}{u_j^{\theta\prime}(C_j^{\theta})}-1\right)+\Lambda^{\theta}-\frac{P_j^{\theta}}{u_j^{\theta\prime}(C_j^{\theta})}\right]u_j^{\theta\prime}(C_j^{\theta})dC_j^{\theta}$$

$$=\mathbb{E}_{\theta}\left[\mathrm{Cov}_{j|\theta}\left(-\frac{P_j^{\theta}}{u_j^{\theta\prime}(C_j^{\theta})},u_j^{\theta\prime}(C_j^{\theta})dC_j^{\theta}\right)\right]+\mathrm{Cov}_{\theta}\left(\Lambda^{\theta}-\mathbb{E}_{j|\theta}\left[\frac{P_j^{\theta}}{u_j^{\theta\prime}(C_j^{\theta})}\right],\mathbb{E}_{j|\theta}\left[u_j^{\theta\prime}(C_j^{\theta})dC_j^{\theta}\right]\right)$$

$$+\mathbb{E}_{\theta}\left[\Lambda^{\theta}\mathbb{E}_{j|\theta}\left[\left(\frac{1}{\mu_j^{\theta}}\frac{\partial\mathcal{U}^{\mathcal{SP},\theta}}{\partial C_j^{\theta}}-u_j^{\theta\prime}(C_j^{\theta})\right)dC_j^{\theta}\right]\right]. \tag{D.31}$$

Consequently, Proposition 1 comes down to

$$dW=\Omega_T+\Omega_{MU}+\Omega_{FE}+\Omega_{TE}+\Omega_R+\Omega_{PM}, \tag{D.32}$$

where

$$\Omega_{PM}=\mathbb{E}_{\theta}\left[\Lambda^{\theta}\mathbb{E}_{j|\theta}\left[\left(\frac{1}{\mu_j^{\theta}}\frac{\partial\mathcal{U}^{\mathcal{SP},\theta}}{\partial C_j^{\theta}}-u_j^{\theta\prime}(C_j)\right)dC_j^{\theta}\right]\right]+\mathbb{E}_{\theta}\left[\Lambda^{\theta}\mathbb{E}_{j|\theta}\left[\frac{\partial\mathcal{U}^{\mathcal{SP},\theta}}{\partial\mu_j^{\theta}}d\ln l_j^{\theta}\right]\right], \tag{D.33}$$

which captures the potential misalignment between the social Planner's welfare assessment of the marginal value of consumption with that of private agents (marginal utility).

Such an approach is useful also in considering welfare criteria and optimal policies that are exclusively based on subcomponents of our decompositions, as in Dávila and Schaab (2022). Consider the following class of welfare criteria:

$$\mathcal{U}^{SP,\theta}(\{C_j^{\theta},\mu_j^{\theta}\})=\sum_{j}(\mu_j^{\theta}+\omega_j^{\theta})u_j^{\theta}(C_j^{\theta})-\psi^{\theta}(\{\mu_j^{\theta}\}). \tag{D.34}$$

Appropriate choice of the type-location specific weights ω_j^θ leads to the following result.

Proposition D.1. *Consider welfare criteria based on* (46) *and* (D.34).

1. *If* $\omega_j^\theta = \frac{\mu_j^\theta}{\Lambda^\theta}\left[\left(\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)} - 1\right) - \left(\Lambda^\theta - \mathbb{E}_{j|\theta}\left[\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}\right]\right)\right]$, *then the decomposition of aggregate welfare in Proposition 1 only consists of* Ω_T, Ω_{FE}, *and* Ω_{TE}.
2. *If* $\omega_j^\theta = -\frac{\mu_j^\theta}{\Lambda^\theta}\left(\Lambda^\theta - \mathbb{E}_{j|\theta}\left[\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}\right]\right)$, *then the decomposition of aggregate welfare in Proposition 1 only consists of* Ω_T, Ω_{MU} Ω_{FE}, *and* Ω_{TE}.
3. *If* $\omega_j^\theta = \frac{\mu_j^\theta}{\Lambda^\theta}\left(\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)} - 1\right)$, *then the decomposition of aggregate welfare in Proposition 1 only consists of* Ω_T, Ω_{FE}, Ω_{TE}, *and* Ω_R.

Proof. As shown earlier, the welfare decomposition with non-welfarist welfare criteria is given by

$$dW = \Omega_T + \Omega_{MU} + \Omega_{FE} + \Omega_{TE} + \Omega_R + \Omega_{PM}, \tag{D.35}$$

where

$$\Omega_{PM} = \mathbb{E}_\theta\left[\Lambda^\theta \mathbb{E}_{j|\theta}\left[\frac{\omega_j^\theta}{\mu_j^\theta} u_j^{\theta\prime}(C_j) dC_j^\theta\right]\right] \tag{D.36}$$

with our assumption (D.34). Note that the second term in Equation (D.33) is absent owing to an envelope condition. Suppose that $\omega_j^\theta = \frac{\mu_j^\theta}{\Lambda^\theta}\left[\left(\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)} - 1\right) - \left(\Lambda^\theta - \mathbb{E}_{j|\theta}\left[\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}\right]\right)\right]$. Then

$$\begin{aligned}
\Omega_{MU} + \Omega_R + \Omega_{PM} = {} & \mathbb{E}_\theta\left[\mathrm{Cov}_{j|\theta}\left(-\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}, u_j^{\theta\prime}(C_j^\theta) dC_j^\theta\right)\right] \\
& + \mathrm{Cov}_\theta\left(\Lambda^\theta - \mathbb{E}_{j|\theta}\left[\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}\right], \mathbb{E}_{j|\theta}\left[u'(C_j^\theta) dC_j^\theta\right]\right) \\
& + \mathbb{E}_\theta\left[\Lambda^\theta \mathbb{E}_{j|\theta}\left[\frac{\omega_j^\theta}{\mu_j^\theta} u'(C_j^\theta) dC_j^\theta\right]\right] \\
= {} & \mathbb{E}_\theta\left[\mathbb{E}_{j|\theta}\left[\left(1 - \frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}\right) u_j^{\theta\prime}(C_j^\theta) dC_j^\theta\right]\right] \\
& + \mathbb{E}_\theta\left(\left(\Lambda^\theta - \mathbb{E}_{j|\theta}\left[\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}\right]\right) \mathbb{E}_{j|\theta}\left[u'(C_j^\theta) dC_j^\theta\right]\right)
\end{aligned}$$

$$
\begin{aligned}
&- \mathbb{E}_\theta \left[\mathbb{E}_{j|\theta} \left[\left(1 - \frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)} \right) u_j^{\theta\prime}(C_j^\theta) dC_j^\theta \right] \right] \\
&- \mathbb{E}_\theta \left(\left(\Lambda^\theta - \mathbb{E}_{j|\theta} \left[\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)} \right] \right) \mathbb{E}_{j|\theta} \left[u'(C_j^\theta) dC_j^\theta \right] \right) \\
=\ & 0. \qquad \text{(D.37)}
\end{aligned}
$$

This proves the first claim. Likewise, if $\omega_j^\theta = -\frac{\mu_j^\theta}{\Lambda^\theta} \left(\Lambda^\theta - \mathbb{E}_{j|\theta} \left[\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)} \right] \right)$, then $\Omega_R + \Omega_{PM} = 0$, and if $\omega_j^\theta = \frac{\mu_j^\theta}{\Lambda^\theta} \left(\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)} - 1 \right)$, then $\Omega_{MU} + \Omega_{PM} = 0$. □

The first case considers a welfare criterion based entirely on aggregate efficiency considerations. The second and third cases incorporate spatial MU dispersion and redistribution considerations, respectively, as well as aggregate efficiency considerations.

Now we derive optimal policies with welfare criteria in each of the three cases discussed in Proposition D.1. In the first case, since Ω_{MU} and Ω_R cancel with Ω_{PM}, optimal transfer policy must satisfy

$$
0 = - \sum_i \mu_i^\theta T_i^\theta \frac{\partial \ln \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta} + \sum_i \mu_i^\theta \sum_{l,k} p_{il,k} y_{il,k} \frac{1}{l_i^\theta} \gamma_{il,k}^\theta \frac{\partial \ln \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta}. \qquad \text{(D.38)}
$$

We can rewrite the above expression to obtain the optimal spatial policy formula that exclusively targets aggregate efficiency considerations:

$$
0 = - \sum_i \frac{\partial \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta} \left[w_i^\theta - P_i^\theta C_i^\theta + \sum_{l,k} p_{il,k} y_{il,k} \frac{1}{l_i^\theta} \gamma_{il,k}^\theta \right]. \qquad \text{(D.39)}
$$

That is, the left-hand side of our baseline formula in Proposition 2 is modified to be zero.

In the second case, the optimal policy formula is

$$
\mu_j^\theta \left[u_j^{\theta\prime}(C_j^\theta) - P_j^\theta \right] = - \sum_i \frac{\partial \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta} \left[w_i^\theta - P_i^\theta C_i^\theta + \sum_{l,k} p_{il,k} y_{il,k} \frac{1}{l_i^\theta} \gamma_{il,k}^\theta \right], \qquad \text{(D.40)}
$$

which incorporates spatial MU dispersion as well as aggregate efficiency considerations.

In the third case, the optimal policy formula is

$$\left(\Lambda^\theta - \mathbb{E}_{j|\theta}\left[\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}\right]\right)\mu_j^\theta u'(C_j^\theta)dC_j^\theta = -\sum_i \frac{\partial \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta}\left[w_i^\theta - P_i^\theta C_i^\theta + \sum_{l,k} p_{il,k}y_{il,k}\frac{1}{l_i^\theta}\gamma_{il,k}^\theta\right], \tag{D.41}$$

which incorporates redistribution considerations as well as aggregate efficiency considerations.

D.6 General Spillovers

In our main model, we assumed that agglomeration externalities are purely a function of local population size (7). In some contexts, researchers specify that a higher population size in the surrounding regions also generates agglomeration spillovers (e.g., Ahlfeldt et al. 2015). In other contexts, researchers also specify that externalities arise from specific producers' input use (e.g., free entry models with labor fixed cost such as Krugman (1991)) or producers' output (e.g., congestion costs from shipment, as in Allen and Arkolakis (2022)). To capture these general externalities, we extend the spillover function (7) such that

$$\mathcal{A}_{ij,k} = A_{ij,k}g_{ij,k}(\{l_\ell^\theta\}, \{l_{ij,k}^\theta\}, y_{ij,k}), \tag{D.42}$$

where the first argument of $g_{ij,k}(\cdot)$ corresponds to the population size across types and locations, the second argument corresponds to labor inputs in production, and the third argument corresponds to output. We also denote the spillover elasticities such that

$$\gamma_{ij,k}^{P,\ell\theta} = \frac{\partial \ln g_{ij,k}}{\partial \ln l_\ell^\theta}, \qquad \gamma_{ij,k}^{L,\theta} = \frac{\partial \ln g_{ij,k}}{\partial \ln l_{ij,k}^\theta}, \qquad \gamma_{ij,k}^{Y,\theta} = \frac{\partial \ln g_{ij,k}}{\partial \ln y_{ij,k}}. \tag{D.43}$$

Under this extension, the only modification in Proposition 1 is in the (iv) technological externality term, which is modified as:

$$\Omega_{TE} = \sum_{j,l,k} p_{jl,k}y_{jl,k}\left(\sum_{\ell,\theta}\gamma_{\ell l,k}^{P,\ell\theta}d\ln l_j^\theta + \sum_\theta \gamma_{jl,k}^{L,\theta}d\ln l_{ij,k}^\theta + \gamma_{jl,k}^{Y}d\ln y_{ij,k}\right). \tag{D.44}$$

This expression recovers the (iv) technological externality term in Proposition 1 if the spillover function only depends on local population size. The only difference here is that the reallocation of population in surrounding regions and of output may have first-order effects on aggregate welfare through additional technological externalities.

E Test of Pareto Efficient Spatial Transfers

Our formula in Proposition 2 can also be used to assess whether the current transfer scheme admits of a Pareto improvement. Since our formula requires the existence of positive Pareto weights $\tilde{\Lambda}^\theta > 0$, equilibrium allocations that lead to negative inferred Pareto weights are Pareto inefficient.

Corollary 2. *If there exists j and θ such that*

$$\mu_j^\theta P_j^\theta < \sum_i \frac{\partial \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta} \left[-T_i^\theta + \sum_{l,k} p_{il,k} y_{il,k} \frac{1}{l_i^\theta} \gamma_{il,k}^\theta \right], \tag{E.1}$$

then there exists an alternative transfer scheme that Pareto improves the original one.

The underlying idea is the same as Werning (2007) in the context of optimal non-linear income taxation. Importantly, this test of Pareto inefficiency does not require researchers to take a stance on Pareto weights or the marginal utility of consumption in each location. The test only requires a handful of sufficient statistics such as price indices, migration elasticities, and agglomeration elasticities.

F Nonparametric Identification of Location Choice System

In this section, we discuss the conditions under which the location choice system $\{\mu_j(\boldsymbol{C})\}$ is nonparametrically identified. To do so, we build on existing results for the nonparametric identification of discrete choice models (Berry and Haile 2014). We abstract from household types and drop superscript θ.

We start by formalizing our econometric environment. Consider a dataset generated by the model of Section 2. We assume that we observe equilibrium configurations under different sets of fundamentals, indexed by $t = 0, 1, \ldots, \mathcal{T}$. A natural interpretation of t is time, but one could also interpret it as types of individuals or demographic groups. We assume that $\{\{w_{j,t}, l_{j,t}, P_{j,t}, T_{j,t}\}, \Pi_t\}$ are observed to the econometrician, so that consumption $C_{j,t} = (w_{j,t} + T_{j,t} + \Pi_t)/P_{j,t}$ is observed as well.

We specify the utility of residing in location j by $u_j(C_{j,t}, \zeta_{j,t}) + \epsilon_{j,t}$, where $\zeta_{j,t}$ is a scalar variable that is unobserved to the econometrician. The unobserved location heterogeneity $\zeta_{j,t}$ captures amenities that vary over j and t. Analogous to Assumption 1 of Berry and Haile (2014), we assume that $\zeta_{j,t}$ only affects location choices through the utility index $u_j(C_{j,t}, \zeta_{j,t})$, but it does not affect the distribution of $\{\epsilon_{j,t}\}$.

Assumption F.1 (Independence)**.** The distribution function of preference shocks $\epsilon_{j,t}$ is independent of $\{\zeta_{j,t}\}$ and t. That is,

$$\mathbb{P}(\epsilon_{1,t} \leq \bar{\epsilon}_1, \ldots, \epsilon_{N,t} \leq \bar{\epsilon}_N | \{\zeta_{j,t}\}) = H(\bar{\epsilon}_1, \ldots, \bar{\epsilon}_N). \tag{F.1}$$

While this assumption is restrictive, we are not imposing any parametric assumption for the distribution function $H(\cdot)$, allowing for flexible correlation of preference shocks across locations.

For the sake of expositional clarity, we also assume in the main text that the unobserved heterogeneity $\zeta_{j,t}$ enters the utility function as a multiplier of consumption: $u_j(C_{j,t}, \zeta_{j,t}) = \bar{u}_j(\zeta_{j,t} C_{j,t})$. As demonstrated by Berry and Haile (2014), this assumption can be relaxed, but such a relaxation requires more technically involved assumptions.

Importantly, we assume that there are vectors of instruments $\mathbf{z}_t$ that are mean independent of the unobserved component of location choice $\ln \zeta_{j,t}$ for all j and t (Assumption F.2) and that there is sufficient variation in $\mathbf{z}_t$ to induce changes in the consumption vector (Assumption F.3).

Assumption F.2 (Exclusion Restriction)**.** $\mathbb{E}[\ln \zeta_{j,t} | \mathbf{z}_t] = 0 \quad$ for all j, t.

Assumption F.3 (Relevance)**.** For all functions $B(C_{jt})$ with finite expectation, if $\mathbb{E}[B(C_{jt}) | \mathbf{z}_t] = 0$ almost surely, then $B(C_{jt}) = 0$ almost surely.

Assumption F.2 is the standard exclusion restriction. Assumption F.3 requires the completeness of the joint distribution of $\{C_{jt}, \mathbf{z}_t\}$, capturing the idea that the instruments $\mathbf{z}_t$ induce sufficient variation in C_{jt}. Under these assumptions, Berry and Haile (2014) show that the location choice system $\mu_{j,t}(\mathbf{C}_t)$ is identified.

Lemma F.1. (Berry and Haile 2014). Suppose Assumptions F.1, F.2, and F.3 hold. Then the location choice system $\mu_{j,t}(\mathbf{C}_t)$ is identified.

Therefore, the location choice system $\{\mu_{j,t}(\boldsymbol{C}_t)\}$ is, at least in principle, nonparametrically identified. At the same time, the data requirements of the excluded instruments

$\mathbf{z}_t$ (Assumptions F.2 and F.3) are substantial. Importantly, to fully identify the flexible substitution patterns for location choice, we need instruments $\mathbf{z}_t$ that induce independent variation in consumption levels in each location $C_{j,t}$. More fundamentally, we need independent observations of equilibrium configurations across different fundamentals t.

G Details on Application I

G.1 Inferring Technology, Ω_T

In this section, we explain the detailed procedure of how we construct productivity growth at the MSA level. Let GDP_i be nominal GDP of MSA i and Y_i be real GDP of MSA i deflated using the GDP deflator of i: P_i^Y. We apply Collorary 1 in Baqaee and Farhi (2024) to express the first-order changes in real GDP of MSA i:

$$d\ln Y_i = \sum_{j,k} \frac{p_{ij,k}y_{ij,k}}{GDP_i} d\ln \mathcal{A}_{ij,k} + \sum_{\theta} \frac{w_i^\theta l_i^\theta}{GDP_i} d\ln l_i^\theta. \tag{G.1}$$

Furthermore, note that

$$d\ln \mathcal{A}_{ij,k} = d\ln A_{ij,k} + \sum_{\theta} \gamma_{ij,k}^\theta d\ln l_i^\theta. \tag{G.2}$$

Using these expressions and given knowledge of $\{\gamma_{ij,k}^\theta\}$, we construct a measure of technological changes at MSA i as follows

$$\sum_{j,k} p_{ij,k}y_{ij,t} d\ln A_{ij,k} = GDP_i \left[d\ln Y_i - \sum_{\theta} \frac{w_i^\theta l_i^\theta}{GDP_i} d\ln l_i^\theta - \sum_{j,k} \frac{p_{ij,k}y_{ij,k}}{GDP_i} \sum_{\theta} \gamma_{ij,k}^\theta d\ln l_i^\theta \right] \tag{G.3}$$

Summing across all MSAs, we obtain the (i) technology term:

$$\begin{aligned} \Omega_{TE} &= \sum_i \sum_{j,k} p_{ij,k}y_{ij,k} d\ln A_{ij,k} \\ &= \sum_i GDP_i \left[d\ln Y_i - \sum_{\theta} \frac{w_i^\theta l_i^\theta}{GDP_i} d\ln l_i^\theta - \sum_{j,k} \frac{p_{ij,k}y_{ij,k}}{GDP_i} \sum_{\theta} \gamma_{ij,k}^\theta d\ln l_i^\theta \right]. \end{aligned} \tag{G.4}$$

By the definition of real GDP,

$$d \ln Y_i = d \ln GDP_i - d \ln P_i^Y. \tag{G.5}$$

Plugging (G.5) back into (G.4), we have

$$\begin{aligned}\Omega_{TE} = \sum_i \left[dGDP_i - \sum_\theta w_i^\theta l_i^\theta d \ln l_i^\theta - \sum_{j,k} p_{ij,k} y_{ij,k} \sum_\theta \gamma_{ij,k}^\theta d \ln l_i^\theta \right] \\ - \sum_l GDP_l \sum_i \frac{GDP_i}{\sum_j GDP_j} d \ln P_i^Y.\end{aligned} \tag{G.6}$$

Given the lack of producer prices at the MSA level, we measure the last term $\sum_i \frac{GDP_i}{\sum_j GDP_j} d \ln P_i^Y$ using the GDP deflator at the national level. We assume the only factor of production is labor and assume away the presence of input-output linkages. These assumptions imply that the GDP of MSA i equals total pre-tax personal income and that sales of skill group θ equal their pre-tax personal income.

G.2 Estimation of Utility Function Parameters

We will now describe the details of estimating the utility function parameters. Recall that we impose the following parametric assumptions:

$$u_{j,t}^\theta (C_{j,t}^\theta) + \epsilon_{j,t}^\theta = \frac{(C_{j,t}^\theta)^{1-\rho_\theta}}{1-\rho_\theta} + \xi_{j,t}^\theta + e_{j,t}^\theta, \tag{G.7}$$

where $e_{j,t}^\theta$ follows an independent type-I extreme value distribution with shape parameter ν_θ. This results in the following logit location choice system

$$\mu_{j,t}^\theta = \frac{\exp\left(\nu_\theta \left[\frac{(C_{j,t}^\theta)^{1-\rho_\theta}}{1-\rho_\theta} + \xi_{j,t}^\theta\right]\right)}{\sum_i \exp\left(\nu_\theta \left[\frac{(C_{i,t}^\theta)^{1-\rho}}{1-\rho_\theta} + \xi_{i,t}^\theta\right]\right)}. \tag{G.8}$$

Taking log, time-differencing, and differencing out with location 1, we obtain

$$\Delta \ln(\mu_{j,t}^\theta / \mu_{1,t}^\theta) = \nu_\theta \left[\Delta \frac{(C_{j,t}^\theta)^{1-\rho_\theta}}{1-\rho_\theta} - \Delta \frac{(C_{1,t}^\theta)^{1-\rho_\theta}}{1-\rho_\theta} \right] + \nu_\theta \left[\Delta \xi_{j,t}^\theta - \Delta \xi_{1,t}^\theta \right], \tag{G.9}$$

where $\Delta x_t \equiv x_t - x_{t-1}$ denotes the time-difference for any variable x_t. The identification threat in estimating Equation (G.9) is that unobserved location-specific amenity shocks $\Delta\xi^{\theta}_{j,t}$ are correlated with changes in consumption. We therefore need instrumental variables $\boldsymbol{Z}_{j,t}$ that are uncorrelated with the location-specific amenity shocks. Define the structural residual as follows:

$$e^{\theta}_{j,t}(\boldsymbol{\beta}_{\theta}) = \Delta \ln(\mu^{\theta}_{j,t}/\mu^{\theta}_{1,t}) - \nu_{\theta}\left[\Delta\frac{(C^{\theta}_{j,t})^{1-\rho_{\theta}}}{1-\rho_{\theta}} - \Delta\frac{(C^{\theta}_{1,t})^{1-\rho_{\theta}}}{1-\rho_{\theta}}\right], \tag{G.10}$$

where $\boldsymbol{\beta}_{\theta} = (\rho_{\theta}, \nu_{\theta})$. Given a vector of $\boldsymbol{Z}_{j,t}$ that satisfies the following moment conditions:

$$\mathbb{E}\left[\left(\Delta\xi^{\theta}_{j,t} - \Delta\xi^{\theta}_{1,t}\right)\boldsymbol{Z}_{j,t}\right] = 0, \tag{G.11}$$

we construct a consistent GMM-estimator of $(\rho_{\theta}, \nu_{\theta})$ that solves

$$\hat{\boldsymbol{\beta}}_{\theta} = \arg\min_{\boldsymbol{\beta}_{\theta}} \boldsymbol{e}^{\theta}(\boldsymbol{\beta}_{\theta})'\boldsymbol{Z}\Phi\boldsymbol{Z}'\boldsymbol{e}^{\theta}(\boldsymbol{\beta}_{\theta}), \tag{G.12}$$

where Φ is a weighting matrix.

To build instrument variables, we construct a shift-share instrument that interacts local industry composition with the national industry employment growth for each skill type θ, similarly to Diamond (2016). Specifically, we construct the following shift-share instrument:

$$z^{\theta}_{j,t} = \sum_{k} \frac{l^{\theta}_{j,k,t-1}}{\sum_k l^{\theta}_{j,k,t-1}} \Delta \ln l^{\theta}_{-j,k,t}, \tag{G.13}$$

where $l^{\theta}_{j,k,t-1}$ denotes the industry k employment of type θ in location j at time $t-1$, and $\Delta \ln l^{\theta}_{-j,k,t}$ is the national industry employment growth of skill θ excluding location j. We construct the above instrument using 5-year samples from the ACS for years 2010 and 2019. We construct an additional instrumental variable that interacts $z^{\theta}_{j,t}$ with consumption growth: $z^{\theta}_{j,t} \times \Delta d \ln C^{\theta}_{j,t}$. We set the weighting matrix as the identity matrix.

We report standard errors based on the consistent estimator of the asymptotic covariance matrix of the GMM estimator $\hat{V}$:

$$\hat{V} = (\hat{G}'\hat{\Omega}^{-1}\hat{G})', \tag{G.14}$$

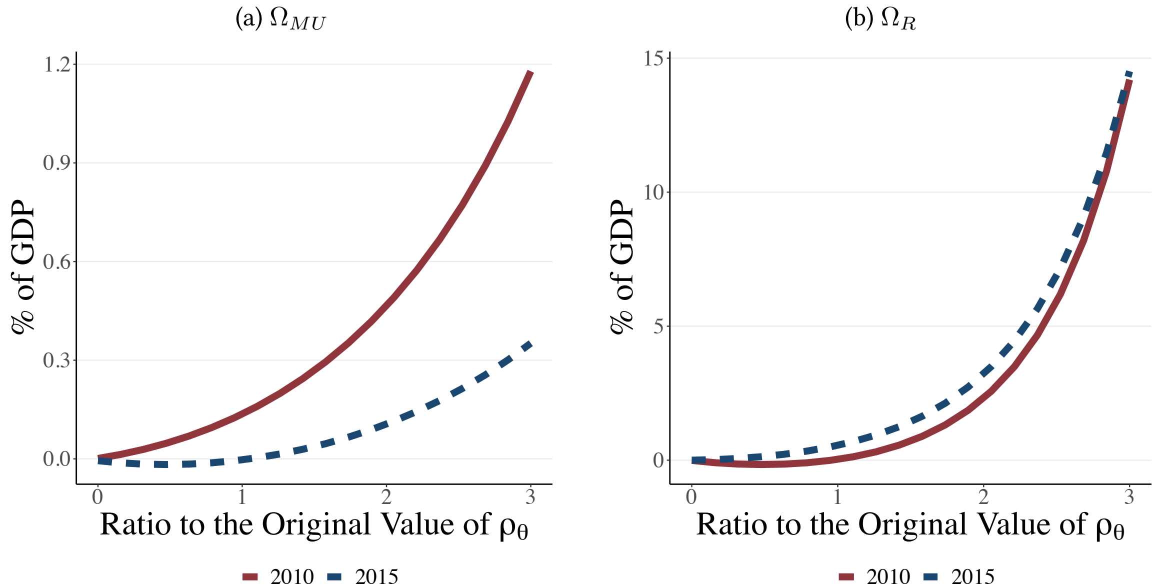

Figure G.1: Sensitivity to the Value of ρ_θ

where

$$\hat{G} \equiv \frac{\partial}{\partial \boldsymbol{\beta}} \left(\boldsymbol{e}^{\theta}(\hat{\boldsymbol{\beta}}_{\theta})' \boldsymbol{Z} \right), \quad \hat{\Omega} = \frac{1}{N} \sum_{i=1}^{N} \left(\boldsymbol{e}_i^{\theta}(\hat{\boldsymbol{\beta}}_{\theta})' \boldsymbol{Z}_i \right) \left(\boldsymbol{e}_i^{\theta}(\hat{\boldsymbol{\beta}}_{\theta})' \boldsymbol{Z}_i \right)'. \tag{G.15}$$

G.3 Sensitivity Analysis

As described in Section 5.1, we consider the sensitivity of our results to our choice of parameter values. We first vary the parameters governing marginal utility: $\{\rho_\theta\}$. We let $\rho_\theta = \bar{\rho}_\theta \times x$, where $\bar{\rho}_\theta$ is our baseline estimate, and vary x. Figure G.1 shows the results. As the value of ρ_θ increases, we see both Ω_{MU} and Ω_R grow substantially in absolute terms.

H Details on Application II

H.1 The Allen and Arkolakis (2022) Model

Allen and Arkolakis (2022) consider an environment with a homogeneous population, so we drop superscript θ. They specify the utility function

$$U_j(C_j, \varepsilon_j) = \ln C_j + \varepsilon_j, \tag{H.1}$$

where ε_j follows type-I extreme value distribution with shape parameter ν.[4] They assume away spatial transfers, so $T_j = 0$ for all j.

The final goods production technology is constant elasticity of substitution (CES), given by

$$C_j = \left(\int_k c_{j,k}^{\frac{\sigma-1}{\sigma}} dk \right)^{\frac{\sigma}{\sigma-1}}, \tag{H.2}$$

where $k \in K = [0, 1]$ indexes the industry, and σ is the elasticity of substitution. The intermediate goods production technology is linear in labor, given by

$$y_{ij,k} = \mathcal{A}_i \tau_{ij,k} l_{ij,k}, \tag{H.3}$$

where $\tau_{ij,k}$ is an iceberg shipment cost, and $\mathcal{A}_i$ is the productivity of region i. Regional productivity is subject to an iso-elastic agglomeration externality in local population size given by

$$\mathcal{A}_i = A_i \left(l_i\right)^{\gamma}, \tag{H.4}$$

where A_i is the fundamental component of productivity.

A key feature of Allen and Arkolakis (2022) is the modeling of the shipment cost $\tau_{ij,k}$ through a route choice problem. Denote by $\mathcal{R}_{ij}$ all possible routes connecting i to j. Formally, $r \in \mathcal{R}_{ij}$ is a sequence of legs (a pair of adjacent locations). Passing through each leg (k, l) incurs iceberg shipment cost t_{kl}. The optimal route choice for producers in region i and sector k implies that the shipment cost $\tau_{ij,k}$ is given by

$$\tau_{ij,k} = \min_{r \in \mathcal{R}_{ij}} \prod_{l=1}^{|r|} t_{r_{l-1} r_l} \epsilon_{ij,k}, \tag{H.5}$$

where $\epsilon_{ij,k}$ is the idiosyncratic cost for each (i, j, k). Finally, they assume that the leg-specific shipment cost may be subject to congestion externalities depending on the traffic passing through the leg. In particular, they assume

$$t_{mn} = \tilde{t}_{mn} \left(\Xi_{mn}\right)^{\lambda}, \tag{H.6}$$

where $\tilde{t}_{mn}$ is the exogenous component of the leg-specific shipment cost (which is in part

[4] As we discuss in Appendix D.4, this specification is isomorphic to assuming an iso-elastic log-linear congestion externality $\nu \ln l_j$ in place of ε_j. Furthermore, as discussed in Section 4.1, welfare changes are invariant by also applying an exponential transformation.

affected by transportation infrastructure), Ξ_{mn} is the value of flows passing through leg (m, n), and λ is the parameter that captures the strength of the congestion externality in shipment costs.

Allen and Arkolakis (2022) use this model to study aggregate welfare changes from a marginal decrease in $\tilde{t}_{mn}$, i.e., a leg-specific improvement in transportation infrastructure. Below, we analyze the same counterfactual experiment, as well as regional productivity changes $d \ln A_i$. It is straightforward to apply our welfare decomposition in Proposition 1. Since we abstract from spatial transfers and multiple types, the (iii) fiscal externality and (v) redistribution terms are zero. The remaining three terms – (i) technology, (ii) MU dispersion, and (iv) technological externality – come down to

$$\Omega_T = -\sum_{k,l} \Xi_{kl} d \ln \tilde{t}_{kl} + \sum_i Y_i d \ln A_i, \tag{H.7}$$

$$\Omega_{MU} = \mathrm{Cov}_j \left(-w_j, d \ln C_j\right), \tag{H.8}$$

$$\Omega_{TE} = \Omega_{TE,S} + \Omega_{TE,A}, \qquad \Omega_{TE,S} = -\lambda \sum_{k,l} \Xi_{kl} d \ln \Xi_{kl}, \quad \Omega_{TE,A} = \gamma \sum_i Y_i d \ln l_i, \tag{H.9}$$

where $\Omega_{TE,S}$ and $\Omega_{TE,A}$ correspond to the technological externalities arising from shipment congestion and productivity agglomeration, respectively.

H.2 Additional Figures

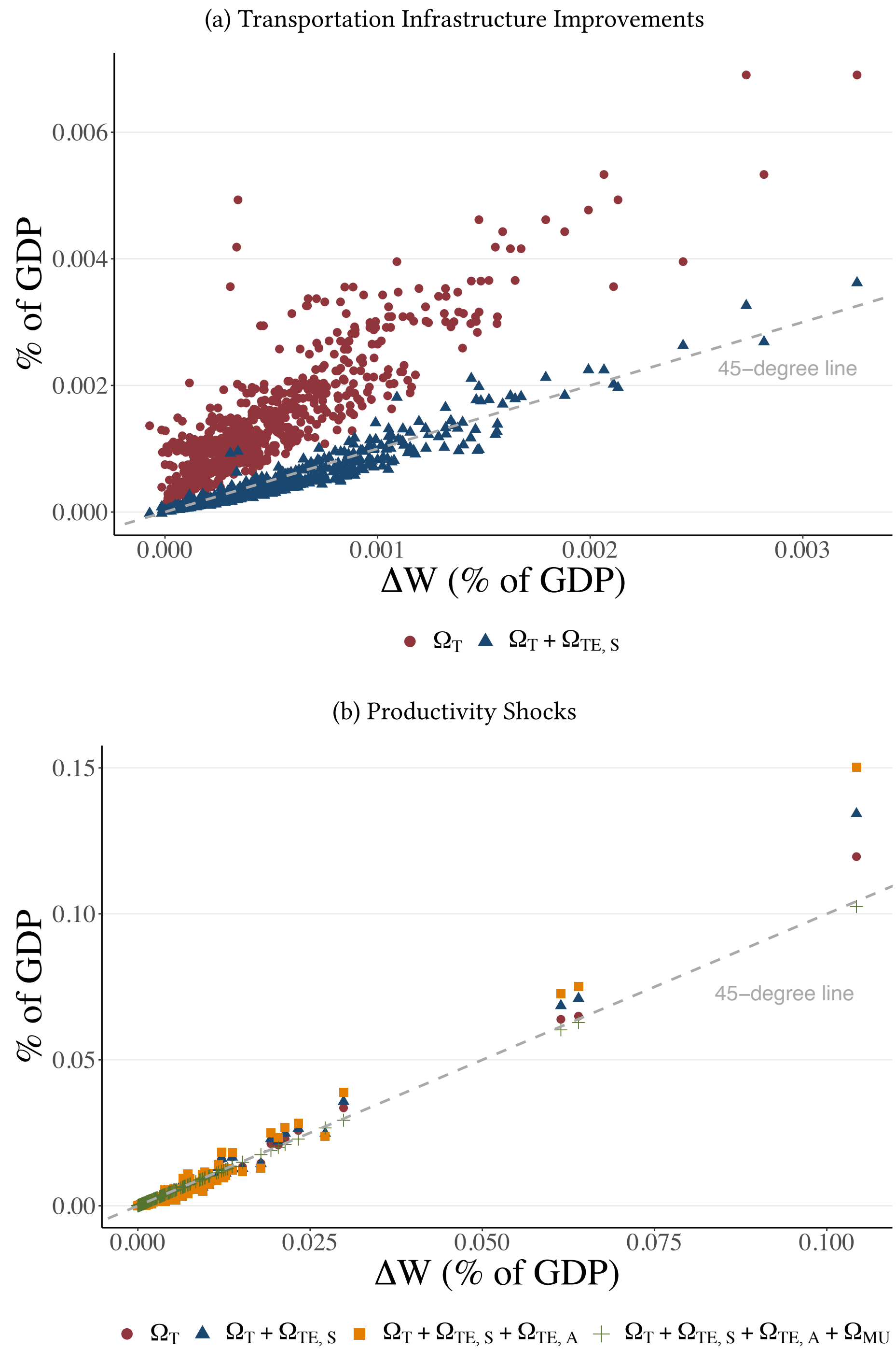

Figure H.1: Welfare Decompositions in the Allen and Arkolakis (2022) Model

Note: This figure plots welfare decomposition described in the main text for each of the counterfactual experiments in Allen and Arkolakis (2022). Panel H.1a is the counterfactual experiment of reducing shipment costs by 1 percent for each of 704 links. Panel H.1b is the counterfactual experiment of increasing productivity by 1 percent for each of 227 CBSAs.