Unpacking Aggregate Welfare in a Spatial Economy [*]

Eric Donald[†] University of Pittsburgh

Masao Fukui[‡] Boston University

Yuhei Miyauchi[§] Boston University

January 7, 2026

Abstract

How do regional productivity shocks or transportation infrastructure improvements affect aggregate welfare? In a general class of spatial equilibrium models, we provide a formula for aggregate welfare changes, decomposed into terms associated with (i) technology (Fogel 1964, Hulten 1978), (ii) spatial dispersion of marginal utility, (iii) fiscal externalities, (iv) technological externalities, and (v) redistribution. We further use this decomposition to derive a general formula for optimal spatial transfers and show that, whenever optimal transfers are in place, the technology term alone captures the aggregate welfare effects of technological shocks. We apply our framework to study welfare gains from improving the US highway network. We find that changes in the spatial dispersion of marginal utility are as important as technological externalities in accounting for the deviations from the Fogel-Hulten benchmark to assess welfare gains.

[*]We thank Rodrigo Adão, Treb Allen, Costas Arkolakis, Yan Bai, Anmol Bhandari, John Sturm Becko, Ariel Burstein, Dani Coen-Pirani, Levi Crews, Dave Donaldson, Maya Eden, Pablo Fajgelbaum, Tishara Garg, Cecile Gaubert, Benny Kleinman, David Lagakos, Ernest Liu, Ezra Oberfield, Andrii Parkhomenko, Natalia Ramondo, Andrés Rodríguez-Clare, Edouard Schaal, Conor Walsh, and Atsushi Yamagishi as well as seminar participants at Barcelona Summer Forum, Chinese University of Hong Kong, Columbia/NYU Spatial-Trade Conference, CRED Workshop on Regional and Urban Economics, Harvard, Midwest Macro Meeting, Minnesota Macro, NBER SI ITI, Philadelphia Fed CURE, Princeton, SED Annual Meeting, SMU, STEG Annual Conference, UC Berkeley, and UEA North American Meeting for their helpful comments. We thank Hannah Rhodenhiser and Maria Mittelbach for excellent research assistance.

[†]University of Pittsburgh (e-mail: eric.donald@pitt.edu).

[‡]Boston University (e-mail: mfukui@bu.edu).

[§]Boston University (e-mail: miyauchi@bu.edu).

1 Introduction

How do regional productivity shocks or transportation infrastructure improvements affect aggregate welfare? To answer these questions, there has been significant progress in the development of quantitative spatial general equilibrium models. These frameworks allow researchers to fit the model to geographically disaggregated data and compute the aggregate welfare implications of a particular shock or policy. While useful, these frameworks are highly complex and parameterized, obscuring which forces, parameters, and data moments govern aggregate welfare effects.

An alternative approach is to apply the macro envelope theorem. Hulten (1978) showed that, in a frictionless representative agent economy, the impact on aggregate GDP of microeconomic TFP shocks is equal to the shocked producer's sales as a share of GDP (i.e. Domar weights). In an evaluation of transportation infrastructure, Fogel (1964) calculated the benefit of shipment cost savings relative to the next best alternative, assuming that reallocation is second order.[1] While these approaches are powerful in their minimal data requirements (Donaldson 2022), they are criticized for ignoring market failures such as agglomeration and congestion externalities.[2] Furthermore, even in the absence of such externalities, it remains unclear whether and how this Fogel–Hulten logic extends to a spatial economy in which heterogeneity in household locations plays a central role.

This paper bridges the gap between these two approaches by providing a theory to unpack aggregate welfare in spatial equilibrium models. We derive a formula for aggregate welfare changes that identifies the precise sources of deviations from the Fogel-Hulten benchmark and link each component to measurable economic objects. Applying our framework to potential improvements to the US highway network, we find that the welfare effects are influenced by forces – particularly the spatial dispersion of marginal utility – that are often overlooked in the application of spatial equilibrium models.

We start by setting up a general class of spatial equilibrium models. The economy consists of a unit mass of households who differ along two dimensions: (i) (exogenous) observable types, such as race or skill level, and (ii) additive idiosyncratic location preferences, capturing factors like proximity to family or personal ties, which are unobservable to researchers and can explain

[1] Fogel (1964) presents two distinct estimates: estimate α, which calculates the value of railroads as the product of shipment value and the reduction in shipment costs, analogous to the approach in Hulten (1978); and estimate β, which additionally accounts for reallocation effects such as changes in land supply. Throughout this paper, references to Fogel (1964) refer specifically to the estimate α, unless otherwise noted.

[2] See Lebergott (1966) and David (1969) for early criticisms of Fogel (1964) due to the omission of agglomeration externalities in his calculation.

why individuals with identical observable traits choose different locations of living. Apart from assuming an additive form for idiosyncratic location preferences, we impose no additional parametric restrictions on these shocks, such as independence across locations or an extreme value distribution.[3] Our framework also nests the case of degenerate idiosyncratic location preferences, following the classic setting studied by Rosen (1979) and Roback (1982). Our framework further accommodates flexible location-specific utility functions as well as local amenities, production functions, input-output linkages, trade frictions, agglomeration and congestion externalities, and government transfers across locations and household types.

We define aggregate welfare via a social welfare function (SWF) that is utilitarian with respect to (potentially degenerate) idiosyncratic location preferences, but which allows arbitrary welfare weights on different household types, following the standard specification in the spatial equilibrium literature (e.g. Roback 1982, Allen and Arkolakis 2014, Redding and Rossi-Hansberg 2017).

Our first theoretical result is a formula for the first-order aggregate welfare impact of technology shocks or transfer policies. We show that it is summarized by five additively separable terms. The first term, (i) Δ technology, is the percentage change in productivity multiplied by the revenue of the region or sector receiving a shock, consistent with the characterization of Fogel (1964) and Hulten (1978). The remaining four terms are (ii) Δ marginal utility (MU) dispersion, (iii) Δ fiscal externalities, (iv) Δ technological externalities, and (v) Δ redistribution, each of which captures a distinct force that generates a deviation from the Fogel-Hulten benchmark.

The second term, (ii) Δ MU dispersion, reflects the reallocation of resources across locations that differ in their marginal utility of income. In spatial equilibrium, agents make location decisions based on utility *levels* (inclusive of idiosyncratic location preferences). This implies that the *marginal* utility of income is not necessarily equalized across locations. As a consequence, a reallocation of consumption across space in response to a shock has a first-order effect on aggregate welfare.

The Δ MU dispersion term is zero if there is no dispersion of marginal utility of income in the pre-shock equilibrium, a knife-edge condition arising under, for example, linear utility *and* equal prices. Interestingly, this term also becomes zero when idiosyncratic location preferences are degenerate, as in the Rosen-Roback framework (Rosen 1979, Roback 1982). This is not because marginal utility is equalized across space; in fact, as originally noted by Mirrlees (1972), dispersion in marginal utility still exists. Rather, it is because, in the absence of idiosyncratic preferences,

[3] We discuss the case of multiplicative idiosyncratic location preferences in Section 4.1.

shocks do not close or widen differences in marginal utility across locations.

The rest of the reallocation terms (iii)-(v) likewise have clear economic interpretations. The third term, (iii) Δ fiscal externality, reflects changes in the government budget induced by population movements in response to shocks. This term is positive if a shock induces reallocation towards locations that are net taxpayers. The fourth term, (iv) Δ technological externality, reflects changes in the agglomeration/congestion externalities induced by population movements or output changes. This term is positive, for example, if a shock induces reallocation towards locations that generate higher technological externalities. The fifth term, (v) Δ redistribution, reflects a reallocation of resources across heterogeneous household types. This term is positive if a shock induces the reallocation of consumption toward household types with higher welfare weights.

Beyond theoretically clarifying the sources of welfare gains and losses, our formula reveals the data moments necessary for measuring aggregate welfare changes. A celebrated feature of Fogel-Hulten's analysis is that it only requires data on changes in productivity and pre-shock equilibrium sales. Our formula sheds light on when and why we require additional information once we depart from the Fogel-Hulten benchmark. Beyond changes in equilibrium consumption and population, three sets of structural parameters and pre-shock equilibrium objects are required for drawing welfare conclusions: the marginal utility of consumption — which governs changes in MU dispersion; agglomeration elasticities — which govern changes in technological externalities; and spatial transfers — which govern changes in fiscal externalities. We argue that all the structural parameters in our welfare formula can be nonparametrically identifiable with suitable exogenous variations, thereby providing a clear mapping from data moments to aggregate welfare changes.

We then use our characterization to derive a nonparametric formula for optimal spatial transfers, which generalizes existing results. Optimality of spatial transfers requires a marginal change in transfers not to improve welfare, so by imposing this requirement in our characterization, we reveal the trade-offs that optimal transfers must navigate. Specifically, our formula trades off the value of closing the spatial dispersion in marginal utility against the fiscal and technological externalities induced by a transfer, echoing the Baily-Chetty formula (Baily 1978, Chetty 2006) in the context of the optimal unemployment insurance literature.

We further show that, if optimal spatial transfers are in place in the pre-shock equilibrium, the first-order welfare effects of technological shocks are fully captured by the (i) Δ technology term alone, thereby recovering the Fogel-Hulten benchmark. This occurs because optimal transfers are designed to neutralize the reallocation terms (ii)–(v), eliminating their first-order impact on aggregate welfare. This result has practical implications for policies that affect productivity, such as investments in transportation infrastructure (e.g. Fajgelbaum and Schaal 2020, Bordeu 2025). Our results suggest that if the government is already implementing optimal spatial transfers, the welfare impact of transportation infrastructure can be evaluated using the Fogel-Hulten benchmark.

A related but distinct question is how spatially disaggregated shocks affect aggregate output. We derive an analogous formula for aggregate output and characterize why it differs from that of aggregate welfare. The formula for aggregate real GDP is similar to that of welfare in that three terms – technology, fiscal externality, and technological externality – enter identically, but differs in that two terms – MU dispersion and redistribution – are now absent. Instead, a new term appears, which reflects the presence of compensating differentials. This term captures that, in spatial equilibrium models, location choices are based on utility and, therefore, do not necessarily maximize aggregate real GDP. The compensating differential term is positive if the population reallocates toward locations with high nominal income, where the compensating differential is low.

In the final part of our paper, we use our framework to unpack the welfare gains from improvements in US highway network. To do so, we first generalize the existing spatial equilibrium framework with route choice and traffic congestion (Allen and Arkolakis 2022) by allowing flexibility in utility functions and spatial transfers, two important welfare-relevant margins highlighted by our formula. We then use this generalized model to quantify the welfare gains from a marginal improvement of each link in the highway network and the precise sources of their deviations from the Fogel-Hulten benchmark.

We find substantial deviations from the Fogel-Hulten benchmark in describing both the average level of and heterogeneity across welfare gains from these link improvements. While our results confirm the importance of technological externalities from traffic congestion, as emphasized by Allen and Arkolakis (2022), we identify a quantitatively comparable role for (ii) Δ marginal utility dispersion in explaining the deviations. In addition, the contribution from (iii) Δ fiscal externalities is meaningful, although to a much lesser extent. Overall, our application sheds light on important sources of welfare gains/losses that are often overlooked in the spatial equilibrium literature.

Related Literature

Our paper contributes to a large and growing literature on spatial equilibrium models (see Redding and Rossi-Hansberg (2017) for a survey) by providing a formula for first-order changes in aggregate welfare relative to the Fogel-Hulten benchmark. A growing body of research uses quantitative spatial equilibrium models to study the aggregate welfare effects of regional productivity shocks (e.g. Caliendo, Parro, Rossi-Hansberg, and Sarte (2018)) or transportation infrastructure (e.g. Allen and Arkolakis (2014, 2022), Donaldson and Hornbeck (2016)). In particular, our paper shares the focus of Zárate (2024), Tsivanidis (2025), and Redding (2025) in studying how welfare gains deviate from the Fogel-Hulten benchmark, or equivalently, the value of travel time saving (VTTS) approach (Small and Verhoef 2007). We provide a general formula that explicitly characterizes the sources of deviation and thereby sheds light on the components that are not typically highlighted in the literature. Our paper is also related to Kleinman, Liu, and Redding (2024) in highlighting the usefulness of first-order approximations in spatial equilibrium models.

In pointing out the spatial dispersion in marginal utility as an important margin of welfare changes, our work builds on Mirrlees (1972), who discusses this issue in a model with stylized geography. In contemporaneous work, Mongey and Waugh (2024) discuss this issue in abstract discrete choice models. They interpret dispersion in marginal utility as a market failure stemming from a lack of insurance over ex-ante uncertain idiosyncratic location preferences. They then characterize the allocation with and without complete insurance markets. The focus of our paper is different. Without taking a strong normative stance on the dispersion of marginal utility as a market failure, we characterize how this feature and other typical considerations in spatial equilibrium models shape aggregate welfare changes beyond Fogel-Hulten's characterization and investigate its connection to optimal spatial transfers.

Our focus on aggregate welfare is related to the work of Bhandari, Evans, Golosov, and Sargent (2023) and Dávila and Schaab (2025a,b), who develop welfare decompositions in general equilibrium models with heterogeneous agents. A key distinction of our approach is that we benchmark welfare changes against Fogel-Hulten's characterization and explicitly decompose the deviation into interpretable and measurable components. This structure enables us to identify the precise sources of deviation from Fogel-Hulten's formula and to link these deviations to the design of optimal spatial transfer policies.

We also contribute to the existing literature on the optimal design of place-based policies. Our optimal spatial transfer formula generalizes existing work focusing on agglomeration externalities (Fajgelbaum and Gaubert 2020, Rossi-Hansberg, Sarte, and Schwartzman 2023) by dispensing with any parametric assumptions for location choice, as well as existing work focusing on redistribution (Davis and Gregory 2021, Ales and Sleet 2022, Gaubert, Kline, Vergara, and Yagan 2025) by introducing a rich production economy, cross-regional trade, and agglomeration externalities. We also contribute to the literature on designing place-based policies that affect productivity, such as transportation infrastructure, by identifying the conditions under which one can rely solely on the Fogel-Hulten benchmark in evaluating those policies.

2 Spatial Equilibrium Framework

This section sets up the general spatial equilibrium framework of our baseline analysis. For expositional clarity, we delegate some additional features (such as shocks to non-market amenities and amenity externalities) to Section 4.

2.1 Setup

There are N locations indexed by $i, j \in \mathcal{N} \equiv \{1, \ldots, N\}$. There is a continuum of households with measure one, indexed by ω. Each household decides its residential location j and the consumption of the location-specific final good aggregator produced using intermediate goods. There are K intermediate goods, some of which can be potentially traded across locations subject to a cost (e.g. food or manufacturing) and some of which are not traded across locations (e.g. housing or nontradable services). Intermediate goods are produced using local labor, intermediate goods, and local fixed factors (e.g. land). Households have ownership of these local fixed factors and earn factor income.

Households are heterogeneous along two dimensions. First, each household ω belongs to an (exogenous) observable type $\theta \in \Theta \equiv \{\theta_1, \ldots, \theta_S\}$, such as race or skill level. We allow for households' preferences for location and consumption, as well as income from labor and fixed factors to flexibly depend on these types. The mass of each type is ℓ^θ with $\sum_\theta \ell^\theta = 1$. Second, within each type θ, households potentially differ in their unobservable preferences over locations, captured by the vector $\boldsymbol{\epsilon}^\theta(\omega) = (\epsilon_1^\theta(\omega), \ldots, \epsilon_N^\theta(\omega))$ for household ω in type θ. These idiosyncratic components reflect unobserved factors – such as proximity to family or personal ties to specific places – that can explain why individuals with identical observable characteristics may make different location choices.

The preference of household ω of type θ residing in location j, and consuming the location-and-type-specific final aggregator C_j^θ, are represented by the following utility function:

$$u_j^\theta(C_j^\theta) + \epsilon_j^\theta(\omega), \tag{1}$$

where $\epsilon_j^\theta(\omega)$ is household ω's idiosyncratic location preference. The common component $u_j^\theta(C_j^\theta)$ can depend on location j and type θ, allowing for amenity differences specific to location and type. We assume that ϵ_j^θ is additively separable. While this assumption is not entirely without loss of generality, it is less restrictive than it may seem, as any monotone transformation of equation (1) preserves identical behavior for location and consumption choices. For instance, a multiplicatively separable specification can be recast into this additive form via a logarithmic transformation.

The budget constraint of a household of type θ residing in location j is

$$P_j^\theta C_j^\theta = w_j^\theta + T_j^\theta + \Pi^\theta, \tag{2}$$

where P_j^θ is the price of final goods, w_j^θ is the wage, and T_j^θ is the net government transfer, all for type θ households in location j. In reality, T_j^θ includes both taxes and transfers explicitly tagged to each location (such as state taxes and transfers in the US) and those set at the national level (such as federal taxes and transfers in the US). We do not impose any additional assumptions about T_j^θ (such as the linearity with respect to nominal wages or income) beyond the government budget constraint (described below) and the assumption that they do not change discontinuously with perturbations of fundamentals. Π^θ is the income from fixed factors for type θ households.

Households choose a location that maximizes their utility. The household ω's optimal location choice conditional on their type θ and location preferences $\boldsymbol{\epsilon}^\theta(\omega)$ solves

$$j^\theta(\omega) \in \arg\max_{i \in \mathcal{N}} u_i^\theta(C_i^\theta) + \epsilon_i^\theta(\omega). \tag{3}$$

We do not make any parametric assumptions for the distribution of $\boldsymbol{\epsilon}^\theta$ beyond the regularity condition that they have a strictly positive density everywhere on $\mathbb{R}^N$ or are degenerate. This specification nests commonly used assumptions about location decisions in the literature. For example, Diamond (2016) considers a case where ϵ_j^θ is distributed according to an i.i.d. type-I extreme value distribution across locations, and McFadden (1978) considers a case where $\boldsymbol{\epsilon}^\theta$ is distributed according to a generalized extreme value distribution with arbitrary correlation across alternatives. Another important case is when ϵ_j^θ is degenerate for all j, as originally considered by Rosen (1979) and Roback (1982). By aggregating across idiosyncratic location preferences, the population size in location j of type θ is given by

$$l_j^\theta = \ell^\theta \mu_j^\theta, \qquad \mu_j^\theta = \int_{\omega \in \mathcal{H}^\theta} \mathbb{I}\left[j = j^\theta(\omega)\right] d\omega, \tag{4}$$

where μ_j^θ is the probability that type θ households choose location j, and $\mathbb{I}\left[j = j^\theta(\omega)\right]$ is an indicator function signifying if household ω chooses location j.

Final goods for type θ households in location j are produced using a constant returns to scale technology over intermediate goods

$$l_j^\theta C_j^\theta = \mathcal{C}_j^\theta(\mathbf{c}_j^\theta), \tag{5}$$

where $\mathbf{c}_j^\theta \equiv \{c_{ij,k}^\theta\}_{i,k}$ denotes a vector of intermediate goods used for final goods production. Here, k indexes intermediate goods and i indexes the origin location of these intermediate goods.

Intermediate good k produced in location i and sold in location j is produced using the following technology

$$y_{ij,k} = \mathcal{A}_{ij,k} f_{ij,k}(\mathbf{l}_{ij,k}, h_{ij,k}, \mathbf{x}_{ij,k}), \tag{6}$$

where $\mathbf{l}_{ij,k} \equiv \{l_{ij,k}^\theta\}_\theta$ denotes labor inputs, $h_{ij,k}$ denotes the local fixed factor input, $\mathcal{A}_{ij,k}$ is Hicks-neutral productivity (including iceberg trade costs), $f_{ij,k}$ is a production function (which we assume to be strictly increasing, concave, differentiable, and constant returns), and $\mathbf{x}_{ij,k} \equiv \{x_{ij,k}^{l,m}\}_{l,m}$ denotes a vector of intermediate inputs, where m indexes the intermediate goods for inputs and l indexes the origin location.[4]

We assume that the supply of the local fixed factor at location j is given exogenously by $\bar{h}_j$. We assume that each type θ household owns α^θ share of fixed factors, where $\sum_\theta \ell^\theta \alpha^\theta = 1$. We also denote the price of the local fixed factor by r_j. Then, the aggregate per-capita return from the fixed factor for a type θ household is given by

$$\Pi^\theta = \alpha^\theta \sum_j r_j \bar{h}_j. \tag{7}$$

[4]Our framework can encompass the case with decreasing returns to scale production function by interpreting the fixed factor $h_{ij,k}$ as a fictitious factor receiving profit.

The government budget constraint is

$$\sum_{\theta}\sum_{j} T_j^{\theta} l_j^{\theta} = 0. \tag{8}$$

Finally, we assume that productivity $\{\mathcal{A}_{ij,k}\}$ is subject to agglomeration spillovers (e.g. productivity increases in the location's population size) or congestion spillovers (e.g. productivity of shipping goods decreases in the volume of shipment). Specifically,[5]

$$\mathcal{A}_{ij,k} = A_{ij,k} g_{ij,k}(\{l_i^{\theta}\}_{\theta}, y_{ij,k}), \tag{9}$$

where $g_{ij,k}(\cdot,\cdot)$ are the spillover functions, and $A_{ij,k}$ is the fundamental component of productivity. Note that we allow for a flexible functional form for spillovers arising from the population size of different household types θ for different locations and goods i, j, k. We denote the elasticity of agglomeration spillovers with respect to population and output as

$$\gamma_{ij,k}^{\theta} \equiv \frac{\partial \ln g_{ij,k}}{\partial \ln l_i^{\theta}}, \qquad \gamma_{ij,k}^{Y} \equiv \frac{\partial \ln g_{ij,k}}{\partial \ln y_{ij,k}}. \tag{10}$$

We define the decentralized equilibrium of this economy as follows.

Definition 1 (Decentralized Equilibrium)**.** A decentralized equilibrium consists of prices $\{\{P_j^{\theta}$, $w_j^{\theta}\}$, $\{p_{ij,k}\}$, $r_j\}$, quantities $\{\{C_j^{\theta}$, $\mathbf{c}_j^{\theta}$, μ_j^{θ}, $l_j^{\theta}\}$, $\{\mathbf{x}_{ij,k}$, $\mathbf{l}_{ij,k}\}\}$, transfers $\{T_j^{\theta}\}$, and productivities $\{\mathcal{A}_{ij,k}\}$ such that:

(i) $\{C_j^{\theta}\}$ satisfies households' budget constraint (2), and $\{\mu_j^{\theta}, l_j^{\theta}\}$ solves households' optimal location choice problem (3) and (4);

(ii) Firms maximize profits

$$\mathbf{c}_j^{\theta} \in \arg\max_{\tilde{\mathbf{c}}_j^{\theta}} P_j^{\theta} \mathcal{C}_j^{\theta}(\tilde{\mathbf{c}}_j^{\theta}) - \sum_{i,k} p_{ij,k} \tilde{c}_{ij,k}^{\theta} \tag{11}$$

and

$$(\mathbf{l}_{ij,k}, h_{ij,k}, \mathbf{x}_{ij,k}) \in \arg\max_{\tilde{\mathbf{l}}_{ij,k}, \tilde{h}_{ij,k}, \tilde{\mathbf{x}}_{ij,k}} p_{ij,k} \mathcal{A}_{ij,k} f_{ij,k}(\tilde{\mathbf{l}}_{ij,k}, \tilde{h}_{ij,k}, \tilde{\mathbf{x}}_{ij,k})$$

[5] By interpreting some intermediate goods k as type θ's labor services, this specification nests general agglomeration spillovers from type θ to another type $\tilde{\theta}$'s labor productivity, nesting the framework of Fajgelbaum and Gaubert (2020). In Section 4.5, we consider the agglomeration externalities that depend not only on local population size but also on that of surrounding regions (e.g. Ahlfeldt, Redding, Sturm, and Wolf 2015) or inputs (e.g. Krugman 1991).

$$
-\sum_{\theta} w_i^{\theta}\tilde{l}^{\theta}_{ij,k} - r_i\tilde{h}_{ij,k} - \sum_{l,m} p_{li,m}\tilde{x}^{l,m}_{ij,k}; \tag{12}
$$

(iii) Goods markets clear

$$
\sum_{\theta} c^{\theta}_{ij,k} + \sum_{l,m} x^{i,k}_{jl,m} = \mathcal{A}_{ij,k} f_{ij,k}(\mathbf{l}_{ij,k}, h_{ij,k}, \mathbf{x}_{ij,k}) \tag{13}
$$

$$
l^{\theta}_j C^{\theta}_j = \mathcal{C}^{\theta}_j(\mathbf{c}^{\theta}_j); \tag{14}
$$

(iv) Labor markets clear

$$
\sum_{i,k} l^{\theta}_{ji,k} = \ell^{\theta}\mu^{\theta}_j; \tag{15}
$$

(iv) Fixed factor markets clear

$$
\sum_{i,k} h_{ji,k} = \bar{h}_j; \tag{16}
$$

(v) Aggregate factor payments Π^{θ} satisfy (7);

(vi) The government budget constraint (8) holds;

(vii) Productivity $\{\mathcal{A}_{ij,k}\}$ is subject to agglomeration or congestion spillovers given by (9).

Throughout the paper, we focus on the case where the decentralized equilibrium is unique and interior ($l^{\theta}_j > 0$ for all j and θ). Since our approach relies on a first-order approximation, this assumption avoids dealing with the case where equilibrium outcomes are non-differentiable with respect to the shock.[6]

2.2 Aggregate Welfare

To define aggregate welfare, we consider a social welfare criterion that is represented by the following social welfare function (SWF):

$$
W = \mathcal{W}(\{W^{\theta}\}), \qquad W^{\theta} \equiv \int_{\omega\in\mathcal{H}^{\theta}} \left(u^{\theta}_j(C^{\theta}_j) + \epsilon^{\theta}_j(\omega)\right) \mathbb{I}\left[j = j^{\theta}(\omega)\right] d\omega. \tag{17}
$$

[6]See Allen and Arkolakis (2014) and Allen, Arkolakis, and Li (2020) for sufficient conditions for equilibrium uniqueness in spatial equilibrium models.

We refer to the welfare weight of type θ households Λ^θ as the marginal value of their expected utility to aggregate welfare, relative to their population size:

$$\Lambda^\theta \equiv \frac{\partial \mathcal{W}(\{W^\theta\})}{\partial W^\theta} \frac{1}{\ell^\theta}. \tag{18}$$

Our social welfare function (17) assumes utilitarianism with respect to (potentially degenerate) additive idiosyncratic location preferences $\epsilon_j^\theta(\omega)$ but allows for arbitrary comparisons across the expected utility of different household types θ.

While this is the standard notion of aggregate welfare used in most applied work (e.g. Roback 1982, Allen and Arkolakis 2014, Redding and Rossi-Hansberg 2017), the assumption of utilitarianism within θ is non-trivial and merits explicit justification. Eden (2020) shows that a utilitarian SWF is equivalent to the following three axioms: (1) independence of irrelevant variations (IIV; i.e. social welfare criteria depend on the distribution of individual welfare gains, not on the idiosyncratic preferences that generate them per se); (2) weak anonymity (i.e. social welfare criteria do not depend on the individual identities of two people with the same preferences); and (3) Pareto condition (i.e. social welfare criteria weakly prefer allocations that all individual weakly prefer).

These axioms clarify the interpretation of our SWF: we treat changes in allocations as normatively equivalent if they yield the same distribution of welfare gains and losses across households within a given type θ, regardless of their idiosyncratic location preferences (IIV) or individual identities (weak anonymity).[7]

While we view our SWF and its underlying axioms as plausible, one may wish to consider alternative aggregate welfare criteria. We revisit this point in Section 3.4.[8]

2.3 Useful Representation Lemmas

We present two lemmas that will be useful later. First, we introduce a convenient alternative representation of location choice decisions. Following Hofbauer and Sandholm (2002), the discrete location choice decision under additive idiosyncratic preferences (1) can be isomorphically

[7] Interestingly, Eden (2020) also shows that no SWF satisfies the three axioms if the utility function cannot be represented with an additive form of idiosyncratic location preferences (1). Intuitively, it is difficult to make interpersonal welfare comparisons if households with different ϵ_j^θ disagree on the welfare gains depending on which location's final goods the gains are evaluated by. See Davis and Gregory (2021) and Gaubert et al. (2025) for a related discussion about the challenges in deriving welfare implications with non-additive specification.

[8] In Section 4.1, we consider a popular alternative approach, in which utility is multiplicatively separable with idiosyncratic location preferences rather than additively separable, and SWF is utilitarian without taking a logarithmic transformation. We show that this specification delivers the identical welfare implications as the additively separable counterpart, provided that preference shocks follow a multivariate Frechét distribution.

represented by households jointly choosing the population size across locations subject to a cost function, as summarized in the following lemma:

Lemma 1 (Hofbauer and Sandholm 2002)**.** *The share of type θ households living in each location $\{\mu_j^\theta\}_j$ can be represented as the solution to the following problem given a vector of equilibrium consumption $\{C_j^\theta\}_j$:*

$$\begin{aligned} W^\theta = &\max_{\{\mu_j^\theta\}_j} \sum_j \mu_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\mu_j^\theta\}) \\ &\textit{s.t.} \quad \sum_j \mu_j^\theta = 1 \end{aligned} \tag{19}$$

for some function $\psi^\theta(\{\mu_j^\theta\})$, which we provide an explicit expression for in Appendix A.1. Moreover, W^θ coincides with the expected utility with respect to $\{\epsilon_j^\theta\}_j$ given optimal location choice in equation (17)*, i.e. $W^\theta = \mathbb{E}[\max_j\{u_j^\theta(C_j^\theta) + \epsilon_j^\theta\}]$.*

The detailed proofs of this lemma and the subsequent propositions of this paper are found in Appendix A. Importantly, $\psi^\theta(\{\mu_j^\theta\})$ summarizes the influence of idiosyncratic location preferences on households' location decisions. If the idiosyncratic location preferences are degenerate (Rosen 1979, Roback 1982), we have $\psi^\theta(\{\mu_j^\theta\}) = 0$. If they follow an i.i.d. type-I extreme value distribution with scale parameter ν, then $\psi^\theta(\{\mu_j^\theta\}) = \frac{1}{\nu}\sum_j \mu_j^\theta \ln \mu_j^\theta$ (Anderson, De Palma, and Thisse 1988). If they follow a type-I generalized extreme value (GEV) with arbitrary correlations (McFadden 1978), $\psi^\theta(\{\mu_j^\theta\}) = \frac{1}{\nu}\sum_j \mu_j^\theta \ln S_j^\theta(\{\mu_i^\theta\})$, where function $S_j^\theta(\cdot)$ depends on the correlation function of $\{\epsilon_j^\theta\}_j$ across alternatives j (see Appendix C).[9]

This lemma is particularly useful because it yields a simple and intuitive expression for how changes in consumption affect expected utility: $\partial W^\theta / \partial C_j^\theta = \mu_j^\theta u_j^{\theta\prime}(C_j^\theta)$. This result follows from the envelope theorem, which implies that the indirect effects of changes in the location choice probabilities μ_j^θ induced by the changes in C_j^θ cancel out to a first order.[10] This result is particularly useful for our analysis of the first-order changes in aggregate welfare.

[9]An alternative interpretation of $\psi^\theta(\cdot)$ is that it captures congestion externalities. For example, the model with idiosyncratic location preferences following an i.i.d. type-I extreme value distribution with scale parameter ν is isomorphic to a model without idiosyncratic location preferences where utility is given by $u_j^\theta(C_j^\theta) - \frac{1}{\nu}\ln \mu_j^\theta$. See Appendix E.3 for further discussion about this isomorphism.

[10]In the original formulation with heterogeneous agents with idiosyncratic preferences $\{\epsilon_j^\theta\}_j$, this result reflects the fact that inframarginal agents (those who strictly prefer location j in baseline equilibrium) experience a utility gain of $u_j^{\theta\prime}(C_j^\theta)dC_j^\theta$, while marginal agents (those who switch to location j from somewhere else) are, by construction, indifferent between location j and their original locations, and therefore do not contribute to first-order welfare changes.

Second, the following lemma shows that the decentralized equilibrium allocation can be represented as the solution to a "pseudo-planning" problem.

Lemma 2. *Any decentralized equilibrium allocation* $\{\{\check{C}_j^\theta, \check{\mathbf{c}}_j^\theta, \check{\mu}_j^\theta, \check{l}_j^\theta\}, \{\check{\mathbf{x}}_{ij,k}, \check{\mathbf{l}}_{ij,k}, \check{h}_{ij,k}, \check{\mathcal{A}}_{ij,k}\}\}$ *can be represented as a solution to the following pseudo-planning problem*

$$W = \max_{\{W^\theta, \{C_j^\theta, \mathbf{c}_j^\theta, \mu_j^\theta\}\}, \{\mathbf{x}_{ij,k}, \mathbf{l}_{ij,k}, h_{ij,k}, \mathcal{A}_{ij,k}\}} \mathcal{W}(\{W^\theta\}) \tag{20}$$

subject to (13)-(16),

$$W^\theta = \sum_j \mu_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\mu_j^\theta\}) \tag{21}$$

$$\{\mu_j^\theta\}_j \in \underset{\{\tilde{\mu}_j^\theta\}_j : \sum_j \tilde{\mu}_j^\theta = 1}{\arg\max} \sum_j \tilde{\mu}_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\tilde{\mu}_j^\theta\}) \tag{22}$$

$$C_j^\theta = \check{C}_j^\theta \tag{23}$$

$$\mathcal{A}_{ij,k} = \check{\mathcal{A}}_{ij,k} \tag{24}$$

The objective function is aggregate welfare. Constraints (13)-(16) correspond to resource constraints. Constraint (22) imposes that location choices are incentive compatible, (23) restricts consumption to be equal to its equilibrium value, and (24) restricts the productivity to be equal to its equilibrium value.

Lemma 2 clarifies the sources of deviation from Hulten (1978) in the first-order welfare effects of technology shocks in a spatial economy. When constraints (22), (23), and (24) are slack, the planner's problem coincides with the first-best planning problem (Appendix B), in which the Planner maximizes aggregate welfare by choosing consumption and population only subject to resource constraints. In this case, the envelope theorem implies that welfare changes are summarized by sales, as in Fogel-Hulten's characterization. However, these constraints are generally binding, leading to deviations from it.

Notably, the deviation from Fogel-Hulten's characterization may arise even in the absence of technological externalities (24). Since such an economy is Pareto efficient, one may wonder how Hulten's theorem could fail. The reason is that, while Pareto efficiency implies that the equilibrium allocation maximizes *some* social welfare function, it generically will not maximize social welfare functions of the form (17), which postulate utilitarian aggregation over (possibly degenerate) idiosyncratic preferences $\epsilon_j^\theta(\omega)$. More generally, in the presence of household heterogeneity, one cannot invoke Hulten's theorem based on Pareto efficiency alone. Below, we show how these considerations — together with technological externalities — generate systematic deviations from the Fogel–Hulten benchmark.

3 Theoretical Analysis

This section provides our theoretical results on the first-order effects of spatially disaggregated technological shocks on aggregate welfare. Section 3.1 provides a formula for aggregate welfare changes. Section 3.2 discusses the data moments that are required to assess welfare changes. Section 3.3 develops an optimal spatial transfer formula and shows how such transfers influence the impact of technology shocks on aggregate welfare. Sections 3.4 and 3.5 extend the analysis for a more general form of SWF and aggregate output.

3.1 Unpacking Welfare Effects of Disaggregated Shocks

For expositional purposes, we introduce the following expectation and covariance operators. The first set of operators takes the expectation and covariance of statistics associated with location j and type θ, weighted by population size $\ell^\theta \mu_j^\theta$:

$$\mathbb{E}_{j,\theta}[X_j^\theta] \equiv \sum_j \sum_\theta \ell^\theta \mu_j^\theta X_j^\theta, \qquad \mathrm{Cov}_{j,\theta}(X_j^\theta, Y_j^\theta) \equiv \mathbb{E}_{j,\theta}[X_j^\theta Y_j^\theta] - \mathbb{E}_{j,\theta}[X_j^\theta]\mathbb{E}_{j,\theta}[Y_j^\theta]. \tag{25}$$

The second set of operators takes the expectation and covariance of statistics associated with location j for a given type θ household, weighted by location share μ_j^θ:

$$\mathbb{E}_{j|\theta}[X_j^\theta] \equiv \sum_j \mu_j^\theta X_j^\theta, \qquad \mathrm{Cov}_{j|\theta}(X_j^\theta, Y_j^\theta) \equiv \mathbb{E}_{j|\theta}[X_j^\theta Y_j^\theta] - \mathbb{E}_{j|\theta}[X_j^\theta]\mathbb{E}_{j|\theta}[Y_j^\theta]. \tag{26}$$

The third set of operators takes the expectation and covariance of statistics associated with type θ households, weighted by population share ℓ^θ:

$$\mathbb{E}_\theta[X^\theta] \equiv \sum_\theta \ell^\theta X^\theta, \qquad \mathrm{Cov}_\theta(X^\theta, Y^\theta) \equiv \mathbb{E}_\theta[X^\theta Y^\theta] - \mathbb{E}_\theta[X^\theta]\mathbb{E}_\theta[Y^\theta]. \tag{27}$$

We also scale the SWF so that the population-weighted average of welfare weights coincides with the weighted average of the inverse of the marginal utility of income $u_j^{\theta\prime}(C_j^\theta)/P_j^\theta$:

$$\mathbb{E}_\theta\left[\Lambda^\theta\right] = \mathbb{E}_{j,\theta}\left[\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}\right]. \tag{28}$$

Since $u_j^{\theta\prime}(C_j^\theta)/P_j^\theta$ represents the marginal resource cost of raising utility for household type θ in location j, this normalization effectively expresses the aggregate welfare W in units of resource cost.[11]

Now consider small changes in the exogenous components of productivity specific to origin location, destination location, and sector: $\{d\ln A_{ij,k}\}$. These shocks can represent region-sector TFP shocks (e.g. Caliendo et al. 2018) or transportation infrastructure changes (e.g. Allen and Arkolakis 2014, Donaldson and Hornbeck 2016).[12] We also allow for the possibility that the structure of transfers may change simultaneously, denoted by $\{dT_j^\theta\}$, either because of exogenous policy changes or as an endogenous response to the productivity shocks.

By applying the envelope theorem to the pseudo-planning problem of Lemma 2, we obtain the following expression for welfare changes:

Proposition 1. *Consider an arbitrary set of small shocks to the exogenous components of productivity $\{d\ln A_{ij,k}\}$, as well as changes in transfers $\{dT_j^\theta\}$, in a decentralized equilibrium. The first-order impact on welfare can be expressed as*

$$\begin{aligned} dW = & \underbrace{\sum_{i,j,k} p_{ij,k} y_{ij,k} d\ln A_{ij,k}}_{\textit{(i)}\ \Delta\ \textit{Technology}\ (\Omega_T)} + \underbrace{\mathbb{E}_\theta\left[Cov_{j|\theta}\left(-\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}, u_j^{\theta\prime}(C_j^\theta)dC_j^\theta\right)\right]}_{\textit{(ii)}\ \Delta\ \textit{MU Dispersion}\ (\Omega_{MU})} + \underbrace{Cov_{j,\theta}\left(-T_j^\theta, d\ln l_j^\theta\right)}_{\textit{(iii)}\ \Delta\ \textit{Fiscal Externality}\ (\Omega_{FE})} \\ & + \underbrace{Cov_{j,\theta}\left(\sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^\theta}\gamma_{jl,k}^\theta, d\ln l_j^\theta\right) + \sum_{i,j,k} p_{ij,k} y_{ij,k} \gamma_{ij,k}^Y d\ln y_{ij,k}}_{\textit{(iv)}\ \Delta\ \textit{Technological Externality}\ (\Omega_{TE})} \\ & + \underbrace{Cov_\theta\left(\Lambda^\theta - \mathbb{E}_{j|\theta}\left[\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}\right], \mathbb{E}_{j|\theta}\left[u_j^{\theta\prime}(C_j^\theta)dC_j^\theta\right]\right)}_{\textit{(v)}\ \Delta\ \textit{Redistribution}\ (\Omega_R)}. \end{aligned} \tag{29}$$

[11]More formally, dW under the normalization (28) can be interpreted as a form of equivalent variation. To see this, consider a compensation scheme where each household type θ in each location j receives a transfer $d\tilde{T}_j^\theta$ such that all experience the same utility gain, du. That is, $du = d\tilde{T}_j^\theta \times u_j^{\theta\prime}(C_j^\theta)/P_j^\theta$ for all j and θ. The total compensation required is then $\mathbb{E}_{j,\theta}\left[d\tilde{T}_j^\theta\right] = \mathbb{E}_\theta\left[\mathbb{E}_{j|\theta}\left[P_j^\theta/u_j^{\theta\prime}(C_j^\theta)\right]du\right] = \mathbb{E}_\theta\left[\Lambda^\theta du\right] = dW$.

[12]In some context, researchers are interested in shocks to amenities instead of productivity. Our analysis includes those cases by interpreting some intermediate goods as local amenities. From a measurement perspective, applying Proposition 1 requires knowledge of prices of the amenities, which are often unobserved and need to be calibrated or estimated. For example, if transportation infrastructure also brings amenity benefits by shortening commuting time, one can use the value of time for $p_{ij,k}$ and commuting time for $y_{ij,k}$ (i.e. Small and Verhoef 2007). In Section 4.2, we provide an alternative expression for Proposition 1 without using amenity prices.

Below, we explain each term of Proposition 1 and illustrate them through special cases. Table 1 summarizes our formula in those special cases.

Technology, Ω_T. The first term of Proposition 1, which we refer to as (i) Δ technology (Ω_T), captures the effects of productivity changes absent the reallocation of resources. Note that $p_{ij,k}y_{ij,k}$ corresponds to the total sales of intermediate inputs k produced in i and sold in j. The observation that total sales summarize the aggregate effects of a shock reflects the celebrated result of Hulten (1978). If the equilibrium maximizes aggregate welfare W, the first term is sufficient for the welfare consequence of disaggregated shocks, to a first-order. However, this is not true in general, and the remaining reallocation terms may become first order.

MU Dispersion, Ω_{MU}. The second term, which we refer to as (ii) Δ MU (marginal utility) dispersion (Ω_{MU}), captures the fact that shocks reallocate resources across locations that potentially differ in their marginal utility of income. A shock leads to an increase in utility of $u_j^{\theta\prime}(C_j^\theta)dC_j^\theta$ in each location j for type θ households. The covariance is positive if utility changes are higher in locations with a lower inverse marginal utility of income $P_j^\theta/u_j^{\theta\prime}(C_j^\theta)$, which can be interpreted as the resource cost required to increase utility of type θ household in location j by one unit. The expectation $\mathbb{E}_\theta[\cdot]$ takes the weighted average of this covariance across household types θ.

This term is generally non-zero if the marginal utility of income is not equalized across locations, and this is precisely what happens in spatial equilibrium. In equilibrium, agents make location decisions based on utility *levels* (inclusive of idiosyncratic location preferences). This implies that the *marginal* utility of income is not necessarily equalized across locations for household type θ. In fact, marginal utility $u_j^{\theta\prime}(C_j^\theta)$ never shows up in any of the equilibrium conditions. Therefore, redistributing resources toward locations with higher marginal utility of income brings gains for the (utilitarian) SWF.

In certain special cases, MU dispersion is absent in the spatial equilibrium. For example, this case arises under linear utility (i.e. $u_j^\theta(C_j^\theta) = B_j^\theta + C_j^\theta$ for some amenity shifter B_j^θ) *and* no trade frictions such that final good prices P_j^θ are equalized across locations j, as considered by Kline and Moretti (2014). Alternatively, if utility functions are logarithmic $u_j^\theta(C_j^\theta) = B_j^\theta + \ln C_j^\theta$, spatial transfers vary only by household type $T_j^\theta = T^\theta$, *and* nominal wages are equalized across locations $w_j^\theta = w^\theta$, spatial dispersion in the marginal utility of income is absent: $u_j^{\theta\prime}(\frac{w^\theta+T^\theta+\Pi^\theta}{P_j^\theta})/P_j^\theta = \frac{1}{w^\theta+T^\theta+\Pi^\theta}$.

An interesting case is an environment with degenerate idiosyncratic location preferences as in the tradition of Rosen (1979) and Roback (1982). As originally observed by Mirrlees (1972),

Table 1: Formula for Aggregate Welfare Changes in Special Cases

	Ω_T	Ω_{MU}	Ω_{FE}	Ω_{TE}	Ω_R
1. Linear utility and no trade frictions	✓		✓	✓	✓
2. Degenerate idiosyncratic location preference	✓		✓	✓	✓
3. No location-specific transfers	✓	✓		✓	✓
4. No technological externalities	✓	✓	✓		✓
5. Single type	✓	✓	✓	✓	
6. No population mobility	✓			✓	✓
7. If optimal spatial transfers are in place	✓				

Notes: This table presents which terms of our formula in Proposition 1 are generically non-zero in each special case.

even without idiosyncratic location preference, equilibria still involves spatial dispersion in the marginal utility of income. However, the *changes* in MU dispersion are always zero in this case. To see why, notice that utility levels are always equalized across locations, and therefore shocks shift the utility levels in all locations (with positive population) by the same amount, i.e. $u_j^{\theta\prime}(C_j^\theta)dC_j^\theta = dW^\theta$. Consequently, the covariance inside Ω_{MU} is always zero. Put differently, shocks fail to close any gap in marginal utility across locations if the idiosyncratic location preferences are degenerate.

Fiscal Externality, Ω_{FE}. The third term, which we refer to as (iii) Δ fiscal externality (Ω_{FE}), comes from the fact that shocks affect the government's budget. If a shock induces population movement toward a location that pays taxes on net (higher $-T_j^\theta$), this term has a positive effect on welfare.[13] This term is absent whenever there are no location-specific transfers within types θ ($T_j^\theta = T^\theta$ for all j and θ) or the shock does not induce any labor reallocation ($d\ln l_j^\theta = 0$ for all j and θ).

Notice also that this term only depends on the transfers at the baseline equilibrium (T_j^θ), not their changes (dT_j^θ), which can occur either because of exogenous policy changes or as an endogenous response to the productivity shocks to balance the government's budget constraint. It does not imply that these changes do not affect the equilibrium; instead, those changes are encapsulated in changes in consumption or population (dC_j^θ, $d\ln l_j^\theta$). Conditional on knowing those changes, dT_j^θ do not directly influence aggregate welfare changes. In Section 3.3, we use this property to derive optimal transfer policies.

[13] In some existing models, researchers assume that some fraction of fixed factor income is rebated to local households directly (such as through local governments' ownership of local fixed factors), which implies that Π_i^θ depends on i (e.g. Caliendo et al. 2018). In such cases, the fiscal externality term is simply modified to capture these local rebates, i.e. replacing the first term in the covariance of Ω_{FE} by $-(T_j^\theta + \Pi_j^\theta)$.

Technological Externality, Ω_{TE}. The fourth term, which we refer to as (iv) Δ technological externality (Ω_{TE}), captures the changes in agglomeration externalities in productivity, arising either through the reallocation of population or output. If a shock induces the population to move toward a location with a higher net agglomeration externality $\sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^\theta} \gamma_{jl,k}^\theta$, the first term of Ω_{TE} has a positive effect on welfare. Similarly, if a shock induces more production in a sector with a higher net agglomeration externality $p_{ij,k} y_{ij,k} \gamma_{ij,k}^Y$, the second term of Ω_{TE} is higher. Clearly, this term becomes zero if there are no technological externalities in the pre-shock equilibrium, i.e. $\gamma_{ij,k}^\theta = 0$ and $\gamma_{ij,k}^Y = 0$ for all i, j, k, and θ.

Importantly, assuming constant elasticity agglomeration externalities ($\gamma_{ij,k}^\theta = \gamma$) alone does not ensure that the first term in (iv) technological externality term is zero. To see why, observe that in a special case with a single sector, single type, no fixed factor, and no intermediate inputs in production, $\sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^\theta} \gamma_{ij,k}^\theta$ simplifies to $w_j \gamma$. Therefore, reallocating the population toward a location with a higher nominal wage generates positive effects on aggregate welfare, as originally observed by Fajgelbaum and Gaubert (2020).

Redistribution, Ω_R. The fifth term, which we refer to as (v) Δ redistribution (Ω_R), is the covariance between the marginal increase in expected utility of type θ households $\mathbb{E}_{j|\theta}[u_j^{\theta\prime}(C_j^\theta) dC_j^\theta]$ and the utility weight net of expected resource cost on those household types, $\Lambda^\theta - \mathbb{E}_{j|\theta}\left[\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}\right]$. Obviously, with a single household type, the (v) Δ redistribution term becomes zero.

No Population Mobility. Our formula nests the case without population mobility by setting $S = N$ and having type θ_i households always locate themselves in location i: $\mu_i^{\theta_i} = 1$. If the population is immobile ($d \ln l_j^\theta = 0$ for all j and θ), then the (ii) Δ MU dispersion and (iii) Δ fiscal externality become zero. Furthermore, the first term in (iv) Δ technological externality is also zero.

3.2 Identifying Aggregate Welfare Changes from Data

An important aspect of Proposition 1 is that it clarifies the data moments necessary for measuring aggregate welfare changes. A celebrated feature of Fogel-Hulten's analysis is that it only requires the changes in productivity $\{d \ln A_{ij,k}\}$ and pre-shock sales of goods $\{p_{ij,k} y_{ij,k}\}$, evident from our (i) Δ technology term. While it is recognized that welfare analysis generally requires richer data if one deviates from the Fogel-Hulten benchmark, it has remained unclear which specific equilibrium objects and structural parameters are necessary in general spatial equilibrium settings. Our formula makes them precise.

In particular, the reallocation terms (ii)-(v) reveal that welfare changes additionally depend on pre-shock final goods prices $\{P_j^\theta\}$, spatial transfers $\{T_j^\theta\}$, and the changes in consumption and population $\{dC_j^\theta, d\ln l_j^\theta\}$.[14] In terms of the structural objects, we additionally need welfare weights $\{\Lambda^\theta\}$, agglomeration externality elasticities $\{\gamma_{ij,k}^\theta\}$, and the spatial dispersion of marginal utility $\{u_j^{\theta\prime}(C_j^\theta)\}$, evaluated around the baseline equilibrium. Choice of $\{\Lambda^\theta\}$ (or more broadly, the choice of SWF (17)) is a normative consideration that cannot be inferred from agents' equilibrium behavior. We argue below that the remaining two objects can be nonparametrically identified with suitable exogenous variations.[15]

For the agglomeration externality elasticities $\{\gamma_{ij,k}^\theta\}$, identification requires the causal effect of exogenous population changes on productivity. The long-standing literature on agglomeration economies has extensively sought to identify these parameters (e.g. see Melo, Graham, and Noland (2009) for a meta-analysis).

The spatial dispersion of marginal utility can be nonparametrically identified from location choice decisions. Denote the location choice probability by $\{\hat{\mu}_j^\theta(\boldsymbol{C}^\theta)\}$ as a solution to equation (19), where $\boldsymbol{C}^\theta$ is a vector of consumption across locations faced by household type θ. Suppose that we have exogenous variation in the consumption of each location so that we can credibly identify this location choice system (e.g. Berry and Haile (2014) establish nonparametric identification of such discrete choice systems). Then, we can identify the relative marginal utility using Lemma 1, as originally suggested by Allen and Rehbeck (2019):

$$\frac{u_j^{\theta\prime}(C_j^\theta)}{u_i^{\theta\prime}(C_i^\theta)} = \frac{\partial \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta} \Big/ \frac{\partial \hat{\mu}_j^\theta(\boldsymbol{C}^\theta)}{\partial C_i^\theta}. \tag{30}$$

Crucially, marginal utility is nonparametrically identified separately from the additive idiosyncratic location preferences ϵ_j^θ.[16]

While it is reassuring that these structural parameters are in principle nonparametrically identified, implementing nonparametric estimation is unrealistic due to the limited availability of ex-

[14]Existing research indicates that $\{dC_j^\theta\}$ and $\{d\ln l_j^\theta\}$ in response to counterfactual shocks $\{d\ln A_{ij,k}\}$ can be nonparametrically identified. In fact, they are uniquely determined by the factor supply (location choice) system, whose identification is established by Berry and Haile (2014) in general discrete choice models, and the factor demand system, whose identification is established by Adão, Costinot, and Donaldson (2017) in general multi-region trade models.

[15]One can potentially reveal $\{\Lambda^\theta\}$ by assuming that observed policies, e.g. $\{T_j^\theta\}$ in our context, are set optimally. See Adão, Costinot, Donaldson, and Sturm (2023) for such an exercise in the context of import tariffs.

[16]Bhattacharya (2015) provides a related but distinct nonparametric identification result of the welfare changes in discrete choice models. In defining equivalent variation, they formulate that compensation can flexibly depend on idiosyncratic preferences, and hence, marginal utility does not play a role. The above result indicates that nonparametric identification can be achieved even if we relax this stringent compensation scheme.

ogenous variation in most applications.[17] Still, this discussion helps to clarify that the dispersion of marginal utility is inherently related to the location choice system. In our application to the US highway network in Section 5, we follow this intuition to estimate the curvature of the utility function.

3.3 Design of Place-Based Policies

In this section, we show that Proposition 1 has broader implications for the design of optimal spatial transfer policy and how such policies interact with the welfare effects of technology shocks. For expositional simplicity, we focus on the case where the technological externality depends only on local population size but not on output, i.e. $\gamma_{ij,k}^{Y} = 0$. In Appendix A.4.4, we show that the same results continue to hold for the case with $\gamma_{ij,k}^{Y} \neq 0$, provided that the Planner simultaneously implements Pigouvian output taxes to fully correct these externalities.

3.3.1 General Optimal Spatial Transfer Formula

We first use Proposition 1 to derive the general formula for optimal spatial transfers. The optimality of spatial transfers requires that any marginal change in consumption dC_i^θ induced by a policy reform cannot improve welfare. Imposing this requirement in Proposition 1 for the case where all technology shocks are zero implies that $dW = \Omega_{MU} + \Omega_{FE} + \Omega_{TE} + \Omega_R = 0$, where $\Omega_T = 0$. Reformulating this, we have the following formula that optimal spatial transfer must satisfy:

Proposition 2. *Assume $\gamma_{ij,k}^{Y} = 0$ for all i, j, k. Assume also that idiosyncratic location preferences are non-degenerate. The optimal spatial transfers must satisfy*

$$\mu_i^\theta \left[\Lambda^\theta u_i^{\theta\prime}(C_i^\theta) - P_i^\theta\right] = Cov_{j|\theta}\left(T_j^\theta - \frac{1}{l_j^\theta}\sum_{l,k} p_{jl,k} y_{jl,k} \gamma_{jl,k}^\theta, \frac{\partial \ln \mu_j^\theta(\boldsymbol{C}^\theta)}{\partial C_i^\theta}\right) \quad \textit{for all } i, \theta. \tag{31}$$

This proposition reveals the key trade-offs associated with optimal spatial transfer policy.[18] The left-hand side of this expression summarizes the marginal benefit from transferring one unit of consumption to location i for type θ. In particular, if weighted marginal utility $\Lambda^\theta u_i^{\theta\prime}(C_i^\theta)$ is high and the associated price P_i^θ is low in location i relative to other locations, the net benefit of

[17] For example, identifying the factor supply system $\{\partial \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)/\partial C_j^\theta\}$ for all i and j requires a long period of data and exogenous variation in consumption at every location.

[18] Appendix F shows how Proposition 2 can be used to test whether a given transfer scheme can be rationalized by some SWF (17) by assessing whether all welfare weights are nonnegative, following the approach of Werning (2007).

transfers to location i tends to be high. On the right-hand side of this equation, we summarize the marginal cost of this transfer through fiscal and technological externalities. In particular, a unit increase of consumption in location i increases population by $\frac{\partial \ln \hat{\mu}_j^\theta}{\partial C_i^\theta}$ in location j. Notice that this relocation happens in all locations, not only in location i. This population relocation is associated with fiscal externalities T_j^θ and technological externalities $\sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^\theta} \gamma_{jl,k}^\theta$.

The above formula has a strong connection to the optimal unemployment insurance literature (Baily 1978, Chetty 2006). In fact, our formula (31) resembles what is often called the Baily-Chetty formula, which balances the trade-off between closing the gap in marginal utility across employed and unemployed workers against generating fiscal externalities by discouraging job search. Relative to the optimal unemployment insurance formula, our formula differs in that it incorporates many possible location choices and the cost includes technological externalities in addition to fiscal externalities.

Proposition 2 is a strict generalization of Fajgelbaum and Gaubert (2020), who study the same problem in a special case where the idiosyncratic location preferences are degenerate.[19] In particular, if we take the limit of the variance of idiosyncratic location preferences to zero, hence $|\frac{\partial \ln \mu_i^\theta}{\partial C_j^\theta}| \to \infty$, the only way to satisfy equation (31) is to set $T_j^\theta - \frac{1}{l_j^\theta} \sum_{l,k} p_{jl,k} y_{jl,k} \gamma_{jl,k}^\theta = E^\theta$ for some constant E^θ, recovering their formula.[20] Proposition 2 also provides a strict generalization of Ales and Sleet (2022) by introducing a rich production economy, cross-regional trade, and agglomeration externalities. In particular, in the absence of trade costs and agglomeration externalities, P_i^θ drops out of the left-hand side as a contributor to marginal utility dispersion, and agglomeration externalities are zero $\gamma_{jl,k}^\theta = 0$ on the right-hand side, recovering their formula.

3.3.2 Impacts of Technology Shocks when Optimal Spatial Transfers are in Place

We now discuss how the welfare effects of technology shocks interact with the optimal transfer policies. The following proposition shows that, if the government is already implementing the optimal spatial transfers in the pre-shock equilibrium, we can invoke Fogel-Hulten's characterization to assess the marginal effects of technological shocks.

Proposition 3. *Assume $\gamma_{ij,k}^Y = 0$ for all i, j, k. Suppose that transfers $\{T_j^\theta\}$ are set so that Proposi-*

[19] They also consider a case where preferences take the form of $U_j^\theta(C_j^\theta, \epsilon_j^\theta) = \tilde{\epsilon}_j^\theta C_j^\theta$ and $\tilde{\epsilon}_j^\theta$ follows an i.i.d. Frechét distribution. As we show in Section 4.1, this specification is isomorphic to its log-transformation $\ln C_j^\theta + \epsilon_j^\theta$, where ϵ_j^θ follows an i.i.d. type-I extreme value distribution. Consequently, our formula nests this case as well.

[20] See Appendix A.4.3 for a more formal treatment of this limit case.

tion 2 holds in the pre-shock equilibrium. Then, Proposition 1 comes down to

$$dW = \underbrace{\sum_{i,j,k} p_{ij,k} y_{ij,k} d \ln A_{ij,k}}_{\textit{(i) } \Delta \textit{ Technology } (\Omega_T)} . \tag{32}$$

This result follows precisely because the prevailing optimal transfers balance the reallocation terms, $\Omega_{MU} + \Omega_{FE} + \Omega_{TE} + \Omega_R = 0$. As a result, reallocation in response to the shocks jointly has no first-order impact on aggregate welfare. Despite reaching the same conclusion as Hulten (1978), the underlying mechanisms are quite different. In a frictionless representative agent economy, any reallocation of resources has no first-order impact on aggregate welfare. In Proposition 3, the reallocation terms sum to zero, although each term is not necessarily zero. This observation resonates with Costinot and Werning (2018), who study the impact of new technologies with distributional consequences and also find that Hulten's characterization holds under second-best policy.

A corollary of Proposition 3 is that the deviations from the Δ technology term documented in applied work (e.g. Caliendo et al. 2018) can be attributed to the suboptimality of prevailing spatial transfers. Specifically, when shocks reallocate consumption toward locations that receive fewer transfers than those implied by the optimality condition (31), the Δ technology term tends to overstate aggregate welfare gains. More broadly, if spatial transfers are designed to maximize subcomponents of social welfare that correspond to terms (ii)–(v) in Proposition 1, deviations from the Fogel-Hulten benchmark reflect a discrepancy between the social preferences of the designer of the observed transfers and those of the analyst (as in equation (17)). We further elaborate on this interpretation in Section 3.4.

This result also has practical implications for policies that affect productivity, such as investments in transportation infrastructure (e.g. Fajgelbaum and Schaal 2020, Bordeu 2025). Our results suggest that if the government is already implementing optimal spatial transfers, the welfare impact of transportation infrastructure can be evaluated using the Fogel-Hulten benchmark.

3.4 Alternative Social Welfare Functions

Our baseline analysis focuses on a weighted utilitarian SWF as discussed in Section 2.2. In some contexts, researchers may want to consider an alternative welfare criterion. In Appendix D, we consider a general (potentially non-welfarist) SWF of the following form:

$$W = \mathcal{W}\left(\{\mathcal{U}^{SP,\theta}(\{C_j^\theta, \mu_j^\theta\})\}\right), \tag{33}$$

where $\mathcal{U}^{SP,\theta}$ is defined arbitrarily on the distribution of consumption and population of type θ households. Appendix D shows that our formula in Proposition 1 includes one additional term. This term captures the potential misalignment between the social Planner's welfare assessment of the marginal value of consumption and that of private agents.

Such an approach is also useful in considering welfare criteria and optimal policies that are exclusively based on subcomponents of our formula, as in Bhandari et al. (2023) and Dávila and Schaab (2025b). In Appendix D, we explicitly construct SWFs that target individual components of our decomposition and derive the corresponding optimal spatial transfer rules. This analysis further suggests that each term in Proposition 1 can, in principle, be interpreted as capturing the gap between the (potentially non-welfarist) social preferences underlying existing spatial transfers and those of the analyst.

3.5 Aggregate Output

We have focused so far on the aggregate welfare. A related but distinct question is how spatially disaggregated shocks affect aggregate output. In this section, we derive an analogous formula for a measure of aggregate output, rather than aggregate welfare. We characterize the precise deviations between these measures, which facilitate the interpretation of each term in Proposition 1.

We define aggregate real GDP using the standard approach of statistical agencies, using a Divisia index that weights each location's real output by its share of nominal GDP:

$$d\ln Y = \sum_i \frac{GDP_i}{GDP}\left(d\ln GDP_i - d\ln P_i\right), \tag{34}$$

where GDP_i corresponds to aggregate nominal GDP of location i. For notational convenience, we take aggregate nominal GDP as the numeraire, so $\sum_i GDP_i = 1$. The following proposition establishes the first-order effects of shocks or transfers on aggregate real GDP.

Proposition 4. *Consider an arbitrary set of small shocks to the exogenous components of productivity $\{d\ln A_{ij,k}\}$, as well as changes in transfers $\{dT_j^\theta\}$, in a decentralized equilibrium. Taking aggregate nominal GDP as the numeraire, the first-order impact on aggregate real GDP can be ex-*

pressed as

$$d\ln Y = \underbrace{\sum_{i,j,k} p_{ij,k} y_{ij,k} d\ln A_{ij,k}}_{\textit{(i) } \Delta \textit{ Technology } (\Omega_T)} + \underbrace{Cov_{j,\theta}\left(-T_j^\theta, d\ln l_j^\theta\right)}_{\textit{(iii) } \Delta \textit{ Fiscal Externality } (\Omega_{FE})} \tag{35}$$

$$+ \underbrace{Cov_{j,\theta}\left(\sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^\theta} \gamma_{jl,k}^\theta, d\ln l_j^\theta\right) + \sum_{i,j,k} p_{ij,k} y_{ij,k} \gamma_{ij,k}^Y d\ln y_{ij,k}}_{\textit{(iv) } \Delta \textit{ Technological Externality } (\Omega_{TE})} + \underbrace{Cov_{j,\theta}\left(w_j^\theta + T_j^\theta, d\ln l_j^\theta\right)}_{\textit{(vi) } \Delta \textit{ Compensating Differential } (\Omega_{CD})} .$$

Relative to Proposition 1, there are both similarities and differences. Similarly to the case of aggregate welfare, the (i) Δ technology, (iii) Δ fiscal externality, and (iv) Δ technological externality terms influence aggregate output changes in an identical manner. The presence of technological externalities and spatial transfers, both of which are forms of wedges, implies that population reallocation induces first-order effects on aggregate output in addition to changes in technology, as in Baqaee and Farhi (2020).

Unlike the case of aggregate welfare, the two terms, (ii) Δ MU dispersion and (v) Δ redistribution, are absent, and there is a new term, (vi) Δ compensating differential (Ω_{CD}). The absence of MU dispersion and redistribution is intuitive because aggregate output does not take into account households' utility. The new term, (vi) Δ compensating differential, reflects the reallocation of the population across locations that differ in terms of their compensating differential. Specifically, when households choose their location, they consider not only nominal income $w_j^\theta + T_j^\theta$, but also utility differences $u_j^\theta(C_j^\theta)$ (including amenities), idiosyncratic preferences ϵ_j^θ, and price indices P_j^θ. Unless all of these elements are absent, in which case nominal income is equalized across space, the (vi) Δ compensating differential term is non-zero. For example, this term is positive if a shock induces population movements toward a location with low compensating differentials (high nominal income relative to utility levels).[21] Such reallocation is not relevant for welfare but affects aggregate real GDP.

An alternative way to understand the gap between aggregate output and aggregate welfare is to reformulate equation (35) as follows:

$$d\ln Y - \mathrm{Cov}_{j,\theta}\left(w_j^\theta, d\ln l_j^\theta\right) = \Omega_T + \Omega_{TE} \tag{36}$$

where the left-hand side, which captures changes in "multifactor productivity" (Solow 1957), is

[21] The covariance between pre-shock utility levels and the changes in population size, net of moving friction function $\psi^\theta(\{\mu_j^\theta\})$, is zero because of the envelope theorem applied to households' location choice.

the measured change in real GDP minus the contribution of the reallocation of labor ($d \ln l_j^\theta$) across locations with different marginal products (w_j^θ). Relative to the changes in multifactor productivity, the changes in aggregate welfare additionally includes (ii) Δ MU dispersion, (iii) Δ fiscal externalities, and (v) Δ redistribution, which are relevant for welfare but not for productivity.[22]

4 Extensions

We will now discuss the scope of our results and argue that our framework can accommodate further extensions and generalizations of our baseline environment.

4.1 Multiplicative Idiosyncratic Location Preference

In our baseline analysis, we have focused on specifications where idiosyncratic location preferences are additively separable. As noted earlier, this is a relatively mild restriction, since any monotonic transformation of equation (1) preserves the same positive predictions for location and consumption choices. For example, a multiplicatively separable specification can be transformed into the additive form through a logarithmic transformation.

At the same time, when researchers work with multiplicative specifications, they often define a utilitarian SWF (or expected utility) under that specification (e.g. Redding 2016), rather than its additive counterpart. In general, these two SWFs are not isomorphic, as they imply different patterns of marginal utility dispersion. In Appendix E.1, we extend Proposition 1 by relaxing the assumption of additive separability in idiosyncratic location preferences. We show that the proposition continues to hold, except for the term associated with changes in marginal utility dispersion (Ω_{MU}). This term is now influenced by the selection of idiosyncratic location preferences, which affects marginal utility in each location, as we further elaborate below.

However, we show that under a common parametrization in the literature – where idiosyncratic location preferences are multiplicatively separable and follow a max-stable multivariate Fréchet distribution – utilitarian SWFs in this setting are isomorphic to those derived from the additive specification.

For brevity, we assume that there is a single household type and drop the θ superscript.[23]

[22] Baqaee and Burstein (2025) defines a measure of the change in aggregate efficiency, defined by the maximum contraction of the consumption possibility set such that it is possible to keep everyone at least indifferent to their status-quo allocation, and shows that this measure coincides with the multifactor productivity up to a first order.

[23] In Appendix A.4, we prove a generalized version of Proposition 5 to multiple household types, where the isomorphism requires redefining welfare weights across household type θ between the two models.

Formally, preferences for households living in location j are given by

$$\begin{aligned} \tilde{U}_j(C_j, \tilde{\epsilon}_j) &= \tilde{u}_j(C_j)\tilde{\epsilon}_j, \\ \text{with} \qquad \mathbb{P}(\tilde{\epsilon}_1 \leq \bar{\epsilon}_1, \ldots, \tilde{\epsilon}_N \leq \bar{\epsilon}_N) &= \exp(-G(K_1(\bar{\epsilon}_1)^{-\nu}, \ldots, K_N(\bar{\epsilon}_N)^{-\nu})), \end{aligned} \tag{37}$$

where $G(.)$ is a function that is homogeneous of degree one, which we call the "correlation function". The key feature of this specification is the max-stability property, where the distribution of the maximum is Fréchet with shape parameter ν.[24]

To examine the properties of this specification, consider the log transformation of utility: $u_j(C_j) = \ln(\tilde{u}_j(C_j))$ and $\epsilon_j = \ln(\tilde{\epsilon}_j)$. It is straightforward to show that ϵ_j follows a multivariate Gumbel distribution with the same correlation function $G(\cdot)$ such that

$$\begin{aligned} U_j(C_j, \epsilon_j) &= u_j(C_j) + \epsilon_j, \\ \text{with} \quad \mathbb{P}(\epsilon_1 \leq \bar{\epsilon}_1, \ldots, \epsilon_N \leq \bar{\epsilon}_N) &= \exp(-G(K_1(\exp(-\nu\bar{\epsilon}_1)), \ldots, K_N(\exp(-\nu\bar{\epsilon}_N)))). \end{aligned} \tag{38}$$

Since $\ln(\cdot)$ is a monotone transformation, the systems (37) and (38) have isomorphic *positive* predictions. The following proposition shows that these two models also deliver isomorphic *normative* predictions under a utilitarian SWF.

Proposition 5. *Consider the spatial equilibrium with multiplicative Fréchet idiosyncratic location preferences with arbitrary correlation* (37). *Let* $\tilde{W} = \int_{\omega \in \mathcal{H}} \tilde{u}_j(C_j)\tilde{\epsilon}_j(\omega)\mathbb{I}\left[j = j(\omega)\right] d\omega$ *be aggregate welfare in this economy, where* $\tilde{\epsilon}_j(\omega)$ *indicates household* ω*'s idiosyncratic location preference for location* j*. Consider another economy where preferences are a log transformation of the first specification, i.e. the additively separable counterpart* (38)*, and the remaining equilibrium conditions are unchanged. Define aggregate welfare of this transformed economy by equation (17). Then, aggregate welfare, as well as the formula of Proposition 1, are identical in both economies up to a multiplicative constant.*

Proposition 5 establishes that the utilitarian SWF in an economy with multiplicative Fréchet location preferences delivers the same normative implications to its additively separable counterpart. This equivalence arises because of a special property of multivariate Fréchet distributions. To see this, the marginal utility of consumption of households in location j in the system (37) is

[24] For example, Redding (2016) is a special case with an i.i.d. Fréchet distribution, which corresponds to the case with $G(x_1, \ldots, x_N) = \sum_{j=1}^{N} x_j$. More generally, this preference specification delivers a generalized extreme value (GEV) demand system with flexible substitution patterns as introduced by McFadden (1978). See also Lind and Ramondo (2023) for the application of this demand system to Ricardian trade models.

given by

$$\mathcal{MU}_j = \frac{\tilde{u}_j'(C_j)}{\tilde{u}_j(C_j)}\mathbb{E}\left[\tilde{u}_j(C_j)\tilde{\epsilon}_j \middle| j = \arg\max_i \tilde{u}_i(C_i)\tilde{\epsilon}_i\right] = u_j'(C_j)\tilde{W}, \tag{39}$$

where we use the fact that $\tilde{u}_j'(C_j)/\tilde{u}_j(C_j) = u_j'(C_j)$. The second transformation of equation (39) follows from the max-stable property of $\tilde{\epsilon}_j$: that the distribution of the maximum follows the same distribution irrespective of the chosen option (McFadden 1978, Lind and Ramondo 2023).

This discussion also highlights that, outside the case of multivariate Fréchet distributions, researchers should exercise caution when transforming utility functions and the associated SWFs, as such transformations may unintentionally alter normative conclusions.

4.2 Shocks to Amenities 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, rather than 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. Appendix E.2 shows an alternative expression for Proposition 1 without using prices for amenities. There, prices on amenities are simply replaced with the marginal utility of amenities.

4.3 Isomorphism of Amenity Externality with Idiosyncratic Preferences

In the quantitative spatial equilibrium literature, researchers often argue that amenity congestion externalities are isomorphic to idiosyncratic location preference and use these specifications interchangeably. For example, Allen and Arkolakis (2014) claim that isoelastic negative amenity externalities are isomorphic to households drawing i.i.d. idiosyncratic location preferences from an extreme-value distribution, both in terms of positive and normative implications. Based on this reasoning, subsequent literature such as Allen and Arkolakis (2022), which is our main application in Section 5, adopts isoelastic negative amenity externalities as a baseline specification, without taking a strong stance on the underlying interpretation.

At first glance, this normative isomorphism may appear puzzling in light of our welfare decomposition. The two specifications generate identical location-choice systems and identical marginal-utility profiles.[25] However, as explained in Section 4.2, the amenity-externality specifi-

[25] Allen and Arkolakis (2014, 2022) specify the utility function in the multiplicative form. As discussed in Section 4.1, this transformation retains the same profiles of marginal utilities with its log transformation (with degenerate idiosyncratic preferences as a special case of multivariate Fréchet distribution).

cation introduces an additional technological externality term – namely, the covariance between amenity externalities and population changes. How can the normative implications remain isomorphic between these two models?

In Appendix E.3, we show how this apparent inconsistency is resolved. Specifically, when amenity externalities take a GEV form – that is, the resulting location choice system exhibits a GEV system – the technological amenity-externality term vanishes. This is because any population reallocation (induced by technological shocks or transfers) cannot alter the aggregate distortion arising from amenity externalities. As a result, amenity externalities affect aggregate welfare solely through location choices, exactly as in models with idiosyncratic preferences. Outside this class, however, the equivalence no longer holds, since the technological amenity-externality term generally is nonzero.

4.4 Idiosyncratic Productivity Shocks

In the baseline analysis, we have considered idiosyncratic shocks to preferences only. Some existing work (e.g. Bryan and Morten 2019, Coen-Pirani 2025) considers instead idiosyncratic shocks to household productivity. Under some additional assumptions, we can tractably incorporate idiosyncratic productivity shocks in addition to idiosyncratic location preferences.

We consider an environment where households of type θ have idiosyncratic productivities in each location $\boldsymbol{z}^\theta \equiv (z_1^\theta, \ldots, z_N^\theta)$ in addition to idiosyncratic location preference $\boldsymbol{\epsilon}^\theta$. The idiosyncratic productivity shocks determine households' endowment of efficiency units of labor in each location. We impose three additional assumptions relative to the baseline model. First, we restrict attention to the case of log utility: $u_j^\theta(c_j^\theta) = B_j \ln c_j^\theta + D_j$. Second, we assume away the presence of fixed factors. Third, we assume that location-specific transfers follow a specific rule such that they are linear in labor income. With these assumptions, the budget constraint of households living in location j is $P_j^\theta c_j^\theta = (1+\tau_j^\theta) w_j^\theta z_j^\theta$, where τ_j^θ denotes the transfer rate. We let $C_j^\theta \equiv w_j^\theta(1+\tau_j^\theta)/P_j^\theta$ denote the consumption per efficiency unit of labor.

Appendix E.4 shows that, in the above environment, Proposition 1 continues to hold with two modifications. First, C_j^θ now denotes consumption per efficiency unit of labor. Second, the (iii) fiscal externality term (Ω_{FE}) contains an additional term that takes into account the changes in the composition of households with different productivities across locations induced by the shock.

4.5 General Spillovers

In our baseline model, we assumed that agglomeration externalities are purely a function of local population size or producers' output (equation 9). 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)). In Appendix E.5, we generalize our results by allowing for a more flexible specification of the spillover function (9).

4.6 Commuting

Our baseline model assumes that households supply labor at the same location as their residential location. In the urban economics literature, it is often assumed that households make separate decisions about their residential and employment location decisions (e.g. Ahlfeldt et al. 2015, Tsivanidis 2025, Zárate 2024). Our framework can be extended to such a framework by reinterpreting households' location decisions j as a combination of residential locations j_1 and work locations j_2, (j_1, j_2). For example, the utility of agents choosing home location j_1 and work location j_2 is given by $u_{j_1 j_2}(C^{\theta}_{j_1 j_2}) + \epsilon^{\theta}_{j_1 j_2}$, where $\epsilon^{\theta}_{j_1 j_2}$ is home-and-work-specific idiosyncratic location preference.[26] Consequently, Proposition 1 remains unchanged by simply replacing j with (j_1, j_2) combinations.

5 Application: Welfare Gains from US Highway Network

In this section, we apply our framework to welfare gains from the US highway network. Every year, more than 150 billion dollars are spent on building, maintaining, or improving highways. Evaluating the impact of infrastructure improvement is critical to determine whether the gains justify these costs. Furthermore, understanding the heterogeneity in welfare gains from different segments of the network is critical to appropriately targeting future infrastructure investment.

Allen and Arkolakis (2022), henceforth AA, study this question using a general equilibrium spatial model with traffic congestion. They found large and highly variable welfare gains from improving transportation infrastructure across different links in the network. AA emphasize that shipment congestion externalities are a key driver of the heterogeneity in deviations for the

[26]This extension accommodates the specification where households consume different consumption bundles depending on the home-work combination, as studied by Miyauchi, Nakajima, and Redding (2025).

Fogel-Hulten benchmark. They demonstrate their importance through simulations that compare models with and without such externalities.

We revisit this question using the AA model, generalizing it along two dimensions. First, we relax their log-linear (Cobb-Douglas) utility assumption and estimate the curvature of the utility function, which plays a critical role in shaping changes in MU dispersion. Second, we incorporate spatial transfers, a salient feature of the US spatial economy, which shapes the changes in fiscal externalities. We calibrate the extended model and simulate the welfare gains from improving each link in the highway network. We then apply Proposition 1 to decompose both the average of and heterogeneity in welfare gains across space into distinct components, thereby clarifying the underlying sources of deviation from the Fogel-Hulten benchmark.

The aim of these generalizations is not to evaluate how our welfare estimates differ from those in AA. Instead, the purpose is to introduce greater flexibility in the potential deviations from the Fogel-Hulten benchmark. We do so by estimating, rather than assuming, the magnitude of marginal utility dispersion; and by incorporating, rather than omitting, fiscal externalities. While congestion externalities are crucial, as emphasized by AA, we argue that changes in MU dispersion, and to a lesser extent, fiscal externalities, also play an important role in explaining the deviations.

5.1 Model Specification

As discussed above, our specification closely follows AA, with two key differences: the form of the utility function and the incorporation of spatial transfers. Abstracting from household types θ as in AA, we specify the utility function as:

$$u_j(C_j) + \varepsilon_j = \frac{C_j^{1-\rho} - 1}{1-\rho} + B_j + \varepsilon_j, \tag{40}$$

where ε_j follows an i.i.d. Type-I extreme value distribution with scale parameter ν, and B_j denotes an exogenous amenity. The parameter ρ governs the curvature of the utility function and thereby determines the degree of marginal utility dispersion.

The case of log utility ($\rho \to 1$) corresponds to the original specification in AA. While AA adopt a different formulation by including an isoelastic congestion externality in B_j and assuming a degenerate distribution of ε_j, their setup is isomorphic to equation (40), as discussed in Section 4.3. As we show there, the scale parameter ν in our formulation corresponds to the inverse of the amenity externality in their environment, and our welfare decomposition remains unchanged under this alternative formulation.[27]

Furthermore, we allow for transfers across locations T_j, so consumption is given by $C_j = (w_j + T_j)/P_j$. We assume that transfers take the following two-part structure:

$$T_j = \varkappa_j w_j + T^*, \tag{41}$$

where $\varkappa_j$ captures the rate of transfer with respect to nominal labor income, and $T^* = -\sum_j \varkappa_j w_j l_j$ is a lump-sum tax/transfer set to satisfy the government budget constraint ($\sum_j T_j l_j = 0$). We also discuss robustness to alternative transfer rules below.

The remaining model follows AA. The nontradable final goods production technology is constant elasticity of substitution (CES), given by

$$C_j = \left(\int_k c_{j,k}^{\frac{\sigma-1}{\sigma}} dk \right)^{\frac{\sigma}{\sigma-1}}, \tag{42}$$

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

$$y_{ij,k} = \frac{\mathcal{A}_i}{\tau_{ij,k}} l_{ij,k}, \tag{43}$$

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 isoelastic agglomeration externality in local population size, given by

$$\mathcal{A}_i = A_i \left(l_i \right)^{\gamma}, \tag{44}$$

where A_i is the fundamental component of productivity.

A key feature of AA is modeling 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 an iceberg shipment cost t_{kl}. Conditional on region j sourcing goods k from region i, the optimal route choice implies that the shipment cost $\tau_{ij,k}$ is given by

$$\tau_{ij,k} = \min_{r \in \mathcal{R}_{ij}} \prod_{l=1}^{|r|} t_{r_{l-1} r_l} \epsilon_{ij,k}, \tag{45}$$

[27] In addition, AA specify the utility function in multiplicative form with amenity. As discussed in Section 4.1, this transformation is inconsequential as long as idiosyncratic preferences follow an extreme value distribution (with degenerate idiosyncratic preferences as a special case).

where $\epsilon_{ij,k}$ is an idiosyncratic cost for each (i, j, k), which follows an i.i.d. Fréchet distribution with dispersion parameter ϑ. Noting that $\tau_{ij,k}$ also follows a Fréchet distribution, the optimal consumption sourcing decision of region j gives rise to a gravity equation in trade flows, as in Eaton and Kortum (2002).

Finally, we assume that the leg-specific shipment cost may be subject to congestion externalities depending on the traffic passing through that leg, so

$$t_{mn} = \tilde{t}_{mn} \left(\Xi_{mn}\right)^{\lambda}, \tag{46}$$

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

We use this model to study aggregate welfare changes from a marginal decrease in $\tilde{t}_{mn}$, i.e. a link-specific improvement in transportation infrastructure. Applying our formula in Proposition 1, the first four terms – (i) technology, (ii) MU dispersion, (iii) fiscal externality, and (iv) technological externality – come down to

$$\Omega_T = -\sum_{m,n} \Xi_{mn} d\ln \tilde{t}_{mn}, \tag{47}$$

$$\Omega_{MU} = \text{Cov}_j \left(-P_j/u_j'(C_j), u_j'(C_j)dC_j\right), \quad \Omega_{FE} = \text{Cov}_j \left(T_j, d\ln l_j\right), \tag{48}$$

$$\Omega_{TE} = \Omega_{TE,A} + \Omega_{TE,S}, \quad \Omega_{TE,A} = \gamma \sum_j w_j l_j d\ln l_j, \qquad \Omega_{TE,S} = -\lambda \sum_{m,n} \Xi_{mn} d\ln \Xi_{mn}, \tag{49}$$

where $\Omega_{TE,S}$ and $\Omega_{TE,A}$ correspond to the technological externalities arising from shipment congestion and productivity agglomeration, respectively.[28] Since we abstract from household types, the redistribution term (Ω_R) is absent.

5.2 Calibration

We choose the same geographical units as AA. Using the 2012 Highway Performance Monitoring System (HPMS) dataset from the Federal Highway Administration, they create the infrastructure network across core-based statistical areas (CBSAs). The resulting network consists of 228 loca-

[28] To apply Proposition 1, the shipment service over each link (m, n) can be interpreted as a distinct intermediate good. Equation (46) slightly deviates from our baseline model in that the shipment congestion externality depends on the value, not the physical unit of outputs. However, such an extension is trivial following a similar logic as in Section 4.5.

tions and 704 links between adjacent nodes.

To conduct our counterfactual simulation, we begin by calibrating the parameters $\{\gamma, \vartheta, \lambda, \rho, \nu, \varkappa_j\}$. For the structural parameters that also appear in AA, we adopt their baseline values. We set the shipment congestion elasticity at $\lambda = 0.092$, the elasticity of localized agglomeration externality at $\gamma = 0.1$, and the dispersion of idiosyncratic shocks for shipment route choice (also corresponding to the trade elasticity) at $\vartheta = 8$.

We estimate the utility function parameters $\{\rho, \nu\}$ using long-run variation in consumption and population across Metropolitan Statistical Areas (MSAs) between 1980 and 2000, employing Generalized Method of Moments (GMM).[29] To construct the moment conditions, we use a shift-share instrumental variable (IV) that interacts local industry composition in 1980 with national industry employment growth between 1980 and 2000, following the approach of Diamond (2016). Additionally, we include an interaction between this IV and baseline consumption levels to identify variation in marginal utility across different consumption levels. To ensure meaningful variation in baseline consumption, we divide the sample into two groups: high-skill individuals (those with four or more years of college education) and low-skill individuals (those with less than four years of college). We then treat the location choices of these two groups as independent samples in the estimation.

We construct real consumption at the MSA level for the years 1980 and 2000, from nominal pre-tax income, consumer price index (CPI), and tax-and-transfer rates. Nominal pre-tax income at the MSA level is derived from the Decennial Census for both years. Since the CPI is available at the MSA level only after 2008 (from the Bureau of Economic Analysis; BEA), we back-cast CPI for 1980 and 2000 by applying state-level inflation rates from Hazell, Herreno, Nakamura, and Steinsson (2022) to the 2008 MSA-level CPI, assuming uniform inflation across MSAs within each state. To construct post-tax-and-transfer income, we apply state-level average tax and transfer rates from the BEA, again assuming uniformity within each state due to the absence of MSA-level tax payment data before 2008. See Appendix G.2.1 for details on the estimation procedure.

Table 2 shows the estimation results. We find a point estimate of ρ at 1.90 with a standard error of 0.78. Therefore, the point estimate suggests a more concave utility function than the commonly-used log-utility specification, while we cannot reject the null hypothesis of $\rho \to 1$ at the five percent critical value. We also find a point estimate of ν at 2.07 with a standard error of 0.85. Notice that, from equation (40), the elasticity of location choice with respect to consumption

[29] We rely on MSAs rather than CBSAs for estimation purposes, as the former are consistently defined and observed in our Decennial Census data.

Table 2: GMM Estimates of Utility Function Parameters

Parameter	Estimates
ρ	1.90
	(0.78)
ν	2.07
	(0.85)

Note: The table reports estimates of (ρ, ν). Standard errors are in parentheses. Estimates of ν are based on the normalization of average consumption in our sample in 2000 to one.

is given by $\partial \ln l_j / \partial \ln C_j = \nu C_j^{1-\rho}(1 - l_j)$. Given our normalization of average consumption in our sample to one in 2000, this implies that the average migration elasticity with respect to consumption during the sample period is approximately 2.07. This estimate is within the range of existing estimates (e.g. Diamond 2016), as well as the baseline value used by AA.[30] In what follows, we take our point estimates as a baseline and discuss the sensitivity of our results to AA's original parametrization.

We calibrate the rates of spatial transfers $\{\varkappa_j\}$ using the observed pre- and post-tax-and-transfer income in 2012. Specifically, we use 2012 county-level estimates from the BEA, which we aggregate to the CBSA level. For each CBSA, we compute the ratio of net taxes and transfers to pre-tax income. Because this measure may not exactly satisfy the government budget constraint $\sum_j T_j l_j = 0$, we adjust $\varkappa_j$ by adding a uniform national constant to ensure balance (see Appendix G.2.2 for details).

Given the above parameter choices, we solve the counterfactual equilibrium following "exact-hat algebra" approach pioneered by Dekle, Eaton, and Kortum (2007) (see Appendix G.1 for the counterfactual equilibrium system). This requires the baseline values of the equilibrium variables $\{l_i, w_i, \Xi_{ij}, C_i\}$. For $\{l_i, w_i\}$, we use the same values as used in AA. For $\{C_i\}$, we set them as post-tax-and-transfer income divided by the CBSA-level CPI for 2012 reported by the BEA.

Finally, we infer the value of traffic over each link $\{\Xi_{ij}\}$ using the average annual daily traffic (AADT) in 2012. Given that these values are traffic counts, not the values they carry over, we infer Ξ_{ij} under the following two assumptions, which slightly differ from AA. First, while AA assume

[30]Again, with $\rho \to 1$, our utility specification in equation (40) is isomorphic to AA's original specification, which takes a multiplicative function of C_j and an isoelastic negative amenity externality with elasticity $-1/\nu$, with degenerate idiosyncratic preferences (Sections 4.1 and 4.3). Their baseline calibration adopts $\nu = 1/0.3$, which is larger but not statistically significantly different from our point estimates.

traffic value is symmetric in both directions for each leg, we allow for asymmetry to accommodate trade imbalance. Second, while AA assume that the aggregate traffic value is exactly equal to national GDP, we instead calibrate it to the observed counterpart. Specifically, we assume that average model-implied expenditure shares for goods produced in other locations coincide with the observed share of tradables in the consumption basket in the US (27.6 percent; Johnson 2017). See Appendix G.2.3 for further details on this procedure. Through this procedure, we find that the sum of traffic value relative to GDP is $\sum_{i,j:i\neq j} \Xi_{ij}/(\sum_i w_i l_i) = 0.726$, somewhat less than the value of one assumed in AA.

5.3 Results

We undertake a counterfactual simulation of decreasing the exogenous component of shipment cost $\tilde{t}_{ij} = \tilde{t}_{ji}$ of each of the 704 links by one percent. To facilitate the interpretation of the welfare changes, we convert these values to an equivalent uniform labor productivity increase ($d \ln A_i = d \ln A$ for all i). This tells us how much of a uniform increase in labor productivity we would need to achieve the same welfare gains as from the transportation infrastructure improvement.[31] On average, a 1% reduction in transportation cost along a link yields a welfare gain equivalent to a 0.00042% increase in uniform productivity.[32] Below, we decompose the estimated welfare gains using Proposition 1, thereby unpacking the sources of deviation from the Fogel-Hulten benchmark.

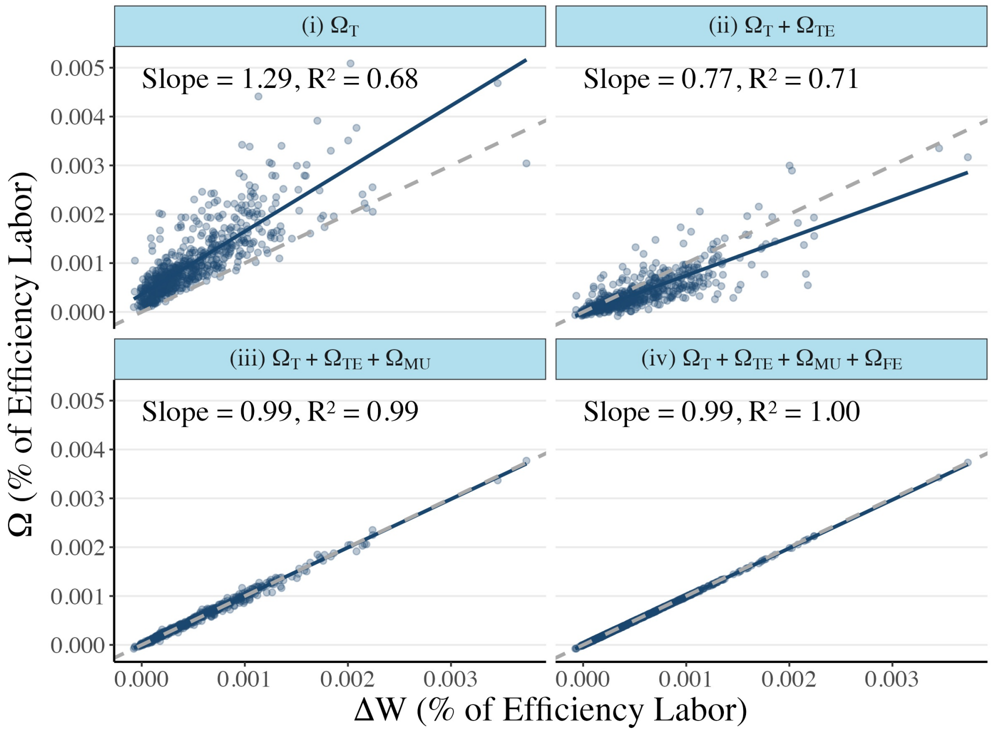
Unpacking the Welfare Gains. Figure 1 displays the estimated welfare gains from improving each link (horizontal axis) against the decomposition terms (vertical axis), cumulatively adding each term one by one over four panels. Each point represents a counterfactual simulation for one of the 704 links in the US highway network. The dotted lines indicate the 45-degree line, and the blue solid lines show the fitted regression lines. Points lying on the 45-degree line imply that the decomposition fully captures the total welfare gains.

Panel (i) plots the Δ technology term (Ω_T) from Proposition 1 against the welfare gain. The Fogel-Hulten benchmark predicts that all dots should lie exactly on the 45-degree line. Instead, we find that the Fogel-Hulten benchmark poorly captures both the average of and heterogeneity in

[31] As noted in the Footnote 11, our normalization allows dW to be interpreted as a form of equivalent variation. Measuring welfare changes in this form (rather than converting them into an equivalent uniform increase in labor productivity, as considered here) yields nearly identical results, with differences of less than one percent.

[32] This average gain is comparable to the results reported in AA, though some differences arise due to variations in model specification and the calibration of traffic values, as discussed earlier.

Figure 1: Unpacking Welfare Gains from Transportation Infrastructure Improvement



Note: This figure plots the welfare gains from 1% reductions in exogenous transportation costs for each of the 704 links in the US highway network. The horizontal axis shows the overall welfare gains (ΔW) from our counterfactual simulation, while the vertical axis shows the sequential decomposition of welfare changes from Proposition 1: (i) Ω_T, (ii) $\Omega_T + \Omega_{TE}$, (iii) $\Omega_T + \Omega_{TE} + \Omega_{MU}$, and (iv) $\Omega_T + \Omega_{TE} + \Omega_{MU} + \Omega_{FE}$. All values are expressed in terms of equivalent uniform labor productivity increases. The final term, $\Omega_T + \Omega_{TE} + \Omega_{MU} + \Omega_{FE}$, does not exactly equal ΔW due to residuals from the first-order approximation. The dotted lines indicate the 45-degree line, and the solid blue lines show the fitted regressions.

welfare gains. First, most data points lie above the 45-degree line, indicating that the technology term tends to overpredict welfare gains on average. In fact, the regression slope of Ω_T on ΔW is 1.29, substantially greater than one. At the same time, the fit is far from perfect. The R^2 from the regression of Ω_T on ΔW is 0.68, suggesting that over 30 percent of the heterogeneity in welfare gains remains unexplained by the technology term alone.

Panel (ii) adds the Δ technological externality term (Ω_{TE}) on top of the Δ technology term. The data points shift closer to the 45-degree line, and the regression slope declines to 0.77, now substantially under-predicting the welfare gains on average. This attenuation of welfare gains reflects the negative impact of shipment congestion externalities on welfare gains. Moreover, the fit remains imperfect, with the R^2 increasing modestly to 0.71 relative to 0.68 in panel (i). AA emphasize the technological externality as a key source of the deviation from the Fogel-Hulten benchmark by comparing counterfactuals with and without such externalities. While we confirm its importance, the technological externality alone is far from enough in guiding the deviations from the Fogel-Hulten benchmark.

Panel (iii) further adds the Δ MU dispersion term (Ω_{MU}). The observations now cluster more tightly around the 45-degree line, with a regression slope of 0.99 and an R^2 of 0.99. While there are some small deviations from the 45-degree line, this suggests that much of the remaining variation is accounted for by the Δ MU dispersion term. This highlights the usefulness of our formula in Proposition 1 in identifying the precise source of deviations from the Fogel-Hulten benchmark that are overlooked in the existing literature.

Panel (iv) further incorporates the Δ fiscal externality term (Ω_{FE}). The slope remains 0.99 and the R^2 improves marginally and is approximately 1.00, with virtually no discernible deviations from the 45-degree line. Therefore, fiscal externality plays a modest role in our welfare decomposition. The fact that four terms jointly capture nearly the entire variation implies that the residual from the first-order approximation in Proposition 1 is negligible. This confirms the tightness and accuracy of the approximation in our empirical context.

Unpacking Deviations from Technology Term. We now take a closer look at each component of the deviation of welfare gains from the Fogel-Hulten benchmark. In the top panel of Figure 2, we plot the distribution of the deviation scaled by the technology term, $(\Delta W - \Omega_T)/\Omega_T$.[33] The average welfare gains from infrastructure improvements are 0.54 log points lower than the prediction based on the technology term, and there is large heterogeneity with a standard deviation of 0.24 log points, consistent with the findings of Panel (i) of Figure 1.

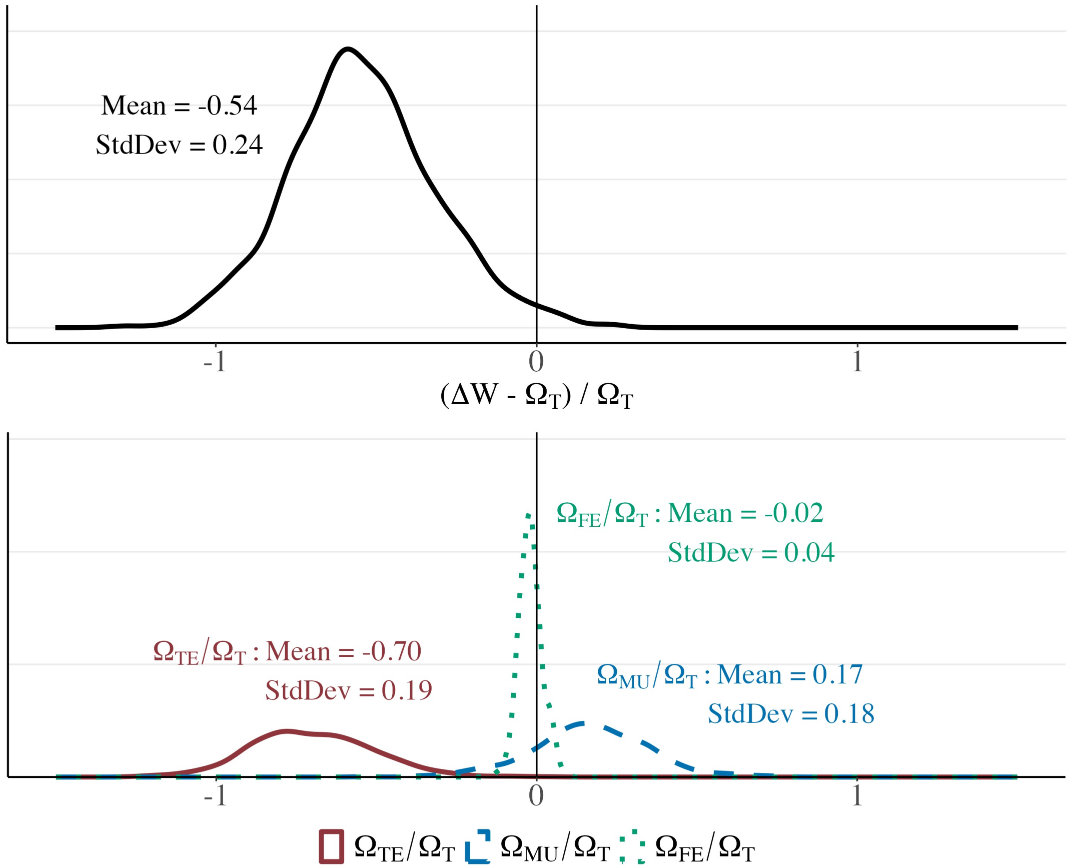
In the bottom panel of Figure 2, we present the three main components of this deviation: Δ technological externality (Ω_{TE}/Ω_T), Δ MU dispersion (Ω_{MU}/Ω_T), and Δ fiscal externality (Ω_{FE}/Ω_T). The technological externality term has a mean of -0.70 log points, reinforcing the interpretation that shipment congestion externalities attenuate the overall welfare gains.[34]

Changes in MU dispersion are also important, with a mean of 0.17 and a standard deviation of 0.18 log points. Positive average gains reflect that infrastructure improvements tend to raise consumption in locations with high marginal utility. The large dispersion across links reflects substantial spatial heterogeneity in the marginal utility of income.

[33] We scale by the technology term Ω_T, not the overall welfare gains ΔW, because the latter sometimes take negative values, while the former always take positive value by construction.

[34] Appendix Table G.3.1 further decomposes the technological externality into contributions from shipment congestion and location-specific agglomeration spillovers. We find that shipment congestion accounts for the bulk of the technological externality ($\Omega_{TE,S}$ in equation 49), with a mean of -0.68 and a standard deviation of 0.19 log points, compared to agglomeration spillovers from local population ($\Omega_{TE,A}$), with a mean of -0.02 log points and a standard deviation of 0.02 log points.

Figure 2: Unpacking Deviations from the Δ Technology Term



Note: The figure presents the density plots of each component of welfare gains in Proposition 1 as a ratio of the Δ technology term (Ω_T) across 704 transportation link improvement simulations.

Finally, the fiscal externality exhibits a more modest but non-negligible deviation, with a mean of -0.02 and a standard deviation of 0.04 log points.

Sensitivity to Alternative Model Specifications. In Appendix G.4, we report the results of the same counterfactual experiment and the associated decomposition of welfare gains under four alternative model specifications: (i) log utility ($\rho \to 1$); (ii) log utility without transfers, which corresponds to AA's original specification; (iii) an alternative transfer rule replacing Specification (41) by lump-sum transfers; (iv) alternative values of the trade elasticity ϑ; and (v) alternative value of the congestion externality elasticity λ. Across all cases, our conclusions remain qualitatively robust: the technology term and MU dispersion term together comprise most of the deviation from the Fogel-Hulten benchmark, with an additional modest contribution from fiscal externalities.

Spatial Patterns of Deviations from Technology Term. To understand the drivers of spatial heterogeneity in deviations from the Fogel-Hulten benchmark, Table 3 presents regressions of these deviations on selected characteristics of each link and its associated nodes. All covariates are standardized to have a unit standard deviation for comparability.

Column (1) shows that deviations from the Fogel-Hulten benchmark are systematically related to the economic characteristics of the connected nodes. Links connecting nodes with one-standard-deviation higher pre-tax income exhibit deviations that are 0.07 log points lower. This pattern is largely driven by Ω_{MU} (Column 2), reflecting lower marginal utility in high-income areas, and is partially offset by $\Omega_{TE,S}$ (Column 4), consistent with denser networks reducing congestion. Links connecting nodes with one-standard-deviation higher net transfers show 0.10 log points lower deviations (Column 1), indicating systematically smaller welfare gains near transfer-receiving locations, and these patterns are primarily driven by Ω_{MU} (Columns 2). In contrast, CPI differences appear only weakly correlated with the deviations, possibly reflecting their limited spatial variation.

Turning to link-specific features, we find that deviations are 0.03 log points higher for links with one-standard-deviation higher traffic, primarily due to $\Omega_{TE,S}$ (Column 4), consistent with the interpretation that nodes with higher traffic tend to have more substitutable routes and therefore tend to generate less congestion. At the same time, in the final row, we do not find a systematically strong relationship of $\Omega_{TE,S}$ with the number of contiguous links, perhaps due to the complex interaction of contiguous links as substitutes (diverting traffic) and complements (attracting additional shipment).

Taking Stock. Overall, these results suggest that welfare gains from transportation infrastructure arise at various margins beyond the Fogel-Hulten benchmark. While the spatial economics literature typically emphasizes technological externalities as the driver of these deviations, we find that changes in MU dispersion, and to a lesser extent fiscal externalities, play quantitatively important roles in understanding both the average and heterogeneity of welfare gains from transportation infrastructure improvements.

6 Concluding Remarks

In a general class of spatial equilibrium models, we have developed a theory to unpack the sources of welfare gains from changes in technology and transfer policies. We provided a formula for

Table 3: Regressions of Deviations from Technology Term on Spatial Economic Characteristics

	$(\Delta W - \Omega_T)$	Ω_{MU}	Ω_{FE}	$\Omega_{TE,S}$	$\Omega_{TE,A}$
	(1)	(2)	(3)	(4)	(5)
log Pre-Tax Income	-0.07	-0.13	-0.004	0.05	0.01
	(0.01)	(0.01)	(0.002)	(0.01)	(0.001)
Net Transfer Rates	-0.10	-0.06	-0.02	-0.02	0.004
	(0.01)	(0.01)	(0.001)	(0.01)	(0.001)
log CPI	-0.002	0.001	0.002	-0.01	0.001
	(0.01)	(0.01)	(0.001)	(0.01)	(0.001)
log Traffic	0.03	-0.002	0.003	0.03	0.0004
	(0.01)	(0.01)	(0.001)	(0.01)	(0.001)
# of Contiguous Links	-0.01	0.003	0.001	-0.01	-0.0003
	(0.01)	(0.01)	(0.001)	(0.01)	(0.001)
Observations	704	704	704	704	704
Adjusted R^2	0.13	0.35	0.31	0.14	0.32

Note: Panel (a) presents the spatial distribution of $(\Delta W - \Omega_T)/\Omega_T$ for each link improvement. The table shows the results from a regression of deviation from technology term $(\Delta W - \Omega_T)$ as well as its decomposition on a set of characteristics of each link that we shock. All the dependent variables are normalized by Ω_T, and all the independent variables are normalized by their standard deviation. All the independent variables, except for log traffic, take the average value between the two nodes that are connected by the link.

welfare changes that characterizes the deviation from the Fogel-Hulten benchmark. We further draw a tight connection between our formula and optimal spatial transfer policies. We used our formula to derive a nonparametric optimal spatial transfer formula, which generalizes those in the existing literature. We then demonstrated that welfare effects can be summarized by the Fogel-Hulten benchmark whenever optimal spatial transfers are in place.

We applied our framework to assess the heterogeneous welfare gains from improvements to transportation links in the US highway network. Our results revealed deviations from the Fogel-Hulten benchmark that are both large on average and heterogeneous across links. While congestion externalities – emphasized in the existing literature – are critical, we found that changes in the dispersion of marginal utility, and to a lesser extent fiscal externalities, also play an important role in explaining these deviations.

Our framework can be useful in a wide range of applications. We further illustrate the use of our framework through two additional applications in the appendix. In Appendix H, we use it to analyze the welfare effects of location-specific productivity shocks using the same model as in Section 5. In contrast to link-specific infrastructure improvements, we find that shipment congestion externalities play a much smaller role, while marginal utility dispersion and fiscal externalities account for a larger share of the welfare impact. This is consistent with the interpretation that city-specific productivity shocks lead to broader shifts in consumption and population, while generating more attenuated and diffused changes in traffic. These findings help clarify that the extent and source of deviation from the Fogel-Hulten benchmark may depend on the nature of the shocks, even when using the same model.

In Appendix I, we apply this framework for an ex-post evaluation of regional productivity growth in the US from 2010-2019. Once changes in consumption and population are observed, our formula requires only a minimal set of parameters, such as marginal utilities and technological externalities, without needing to specify a full structural model. We again find that the Fogel-Hulten benchmark fails to predict the welfare gains from spatially heterogeneous economic growth. MU dispersion and redistribution between high- and low-skill workers (using a utilitarian SWF across types) play an important role in explaining the deviations. In contrast, fiscal and technological externalities play a more modest role during this period.

While our framework is static, many interesting questions in spatial economics relate to the dynamics of economic activity and population mobility. In separate ongoing work (Donald, Fukui, and Miyauchi 2025), we tackle this question by studying optimal transfer policy in a dynamic environment.

References

Adão, R., A. Costinot, and D. Donaldson (2017). Nonparametric Counterfactual Predictions in Neoclassical Models of International Trade. *American Economic Review 107*(3), 633–689.

Adão, R., A. Costinot, D. Donaldson, and J. Sturm (2023). Why is Trade Not Free? A Revealed Preference Approach. Working Paper 31798, National Bureau of Economic Research.

Ahlfeldt, G., S. Redding, D. Sturm, and N. Wolf (2015). The Economics of Density: Evidence from the Berlin Wall. *Econometrica 83*(6), 2127–2189.

Ales, L. and C. Sleet (2022). Optimal Taxation of Income-Generating Choice. *Econometrica 90*(5), 2397–2436.

Allen, R. and J. Rehbeck (2019). Identification with Additively Separable Heterogeneity. *Econometrica 87*(3), 1021–1054.

Allen, T. and C. Arkolakis (2014). Trade and the Topography of the Spatial Economy. *The Quarterly Journal of Economics 129*(3), 1085–1140.

Allen, T. and C. Arkolakis (2022). The Welfare Effects of Transportation Infrastructure Improvements. *The Review of Economic Studies 89*(6), 2911–2957.

Allen, T., C. Arkolakis, and X. Li (2020). On the Equilibrium Properties of Network Models with Heterogeneous Agents. Working Paper 27837, National Bureau of Economic Research.

Anderson, S., A. De Palma, and J.-F. Thisse (1988). A Representative Consumer Theory of the Logit Model. *International Economic Review*, 461–466.

Baily, M. (1978). Some Aspects of Optimal Unemployment Insurance. *Journal of Public Economics 10*(3), 379–402.

Baqaee, D. and A. Burstein (2025). Aggregate Efficiency with Heterogeneous Agents. *Working Paper*.

Baqaee, D. and E. Farhi (2020). Productivity and Misallocation in General Equilibrium. *The Quarterly Journal of Economics 135*(1), 105–163.

Baqaee, D. and E. Farhi (2024). Networks, Barriers, and Trade. *Econometrica 92*(2), 505–541.

Berry, S., A. Gandhi, and P. Haile (2013). Connected Substitutes and Invertibility of Demand. *Econometrica 81*(5), 2087–2111.

Berry, S. and P. Haile (2014). Identification in Differentiated Products Markets Using Market Level Data. *Econometrica 82*(5), 1749–1797.

Bhandari, A., D. Evans, M. Golosov, and T. Sargent (2023). Efficiency, Insurance, and Redistribution Effects of Government Policies. *Working Paper*.

Bhattacharya, D. (2015). Nonparametric Welfare Analysis for Discrete Choice. *Econometrica 83*(2), 617–649.

Bordeu, O. (2025). Commuting Infrastructure in Fragmented Cities. *Working Paper*.

Bryan, G. and M. Morten (2019). The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia. *Journal of Political Economy 127*(5), 2229–2268.

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2018). The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy. *The Review of Economic Studies 85*(4), 2042–2096.

Chetty, R. (2006). A General Formula for the Optimal Level of Social Insurance. *Journal of Public Economics 90*(10-11), 1879–1901.

Coen-Pirani, D. (2025). Tax Progressivity and Mobility Costs. *Regional Science and Urban Economics 114*, 104126.

Costinot, A. and I. Werning (2018). Robots, Trade, and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation. Working Paper 25103, National Bureau of Economic Research.

David, P. (1969). Transport Innovation and Economic Growth: Professor Fogel on and off the Rails. *The Economic History Review 22*(3), 506–525.

Dávila, E. and A. Schaab (2025a). Welfare Accounting. *Working Paper*.

Dávila, E. and A. Schaab (2025b). Welfare Assessments with Heterogeneous Individuals. *Journal of Political Economy 133*(9), 2918–2961.

Davis, M. and J. Gregory (2021). Place-Based Redistribution in Location Choice Models. Working Paper 29045, National Bureau of Economic Research.

Dekle, R., J. Eaton, and S. Kortum (2007). Unbalanced Trade. *American Economic Review 97*(2), 351–355.

Diamond, R. (2016). The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2000. *American Economic Review 106*(3), 479–524.

Donald, E., M. Fukui, and Y. Miyauchi (2025). Optimal Dynamic Spatial Policy. Working Paper 34290, National Bureau of Economic Research.

Donaldson, D. (2022). Blending Theory and Data: A Space Odyssey. *Journal of Economic Perspectives 36*(3), 185–210.

Donaldson, D. and R. Hornbeck (2016). Railroads and American Economic Growth: A "Market Access" Approach. *The Quarterly Journal of Economics 131*(2), 799–858.

Eaton, J. and S. Kortum (2002). Technology, Geography, and Trade. *Econometrica 70*(5), 1741–1779.

Eden, M. (2020). Aggregating Welfare Gains. Discussion Paper DP14783, Centre for Economic Policy Research.

Fajgelbaum, P. and C. Gaubert (2020). Optimal Spatial Policies, Geography, and Sorting. *The Quarterly Journal of Economics 135*(2), 959–1036.

Fajgelbaum, P. and E. Schaal (2020). Optimal Transport Networks in Spatial Equilibrium. *Econometrica 88*(4), 1411–1452.

Flood, S., M. King, R. Rodgers, S. Ruggles, J. R. Warren, and M. Westberry (2023). IPUMS CPS: Version 11.0 [dataset]. *Minneapolis, MN: IPUMS 10*, D030.

Fogel, R. (1964). *Railroads and American Economic Growth.* Johns Hopkins Press Baltimore.

Gaubert, C., P. Kline, D. Vergara, and D. Yagan (2025). Place-based redistribution. *American Economic Review 115*(10), 3415–3450.

Hazell, J., J. Herreno, E. Nakamura, and J. Steinsson (2022). The Slope of the Phillips Curve: Evidence from U.S. States. *The Quarterly Journal of Economics 137*(3), 1299–1344.

Hofbauer, J. and W. Sandholm (2002). On the Global Convergence of Stochastic Fictitious Play. *Econometrica 70*(6), 2265–2294.

Hulten, C. (1978). Growth Accounting with Intermediate Inputs. *The Review of Economic Studies 45*(3), 511–518.

Johnson, N. (2017). Tradable and Nontradable Inflation Indexes: Replicating New Zealand's Tradable Indexes with BLS CPI Data. *Monthly Labor Review*, 1–25.

Kleinman, B., E. Liu, and S. Redding (2024). The Linear Algebra of Economic Geography Models. *AEA Papers and Proceedings 114*, 328–33.

Kline, P. and E. Moretti (2014). People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs. *Annual Review of Economics 6*(1), 629–662.

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

Lebergott, S. (1966). United States Transport Advance and Externalities. *The Journal of Economic History 26*(4), 437–461.

Lind, N. and N. Ramondo (2023). Trade with Correlation. *American Economic Review 113*(2), 317–353.

McFadden, D. (1978). Modeling the Choice of Residential Location. *Transportation Research Record* (673).

Melo, P., D. Graham, and R. Noland (2009). A Meta-Analysis of Estimates of Urban Agglomeration Economies. *Regional Science and Urban Economics 39*(3), 332–342.

Mirrlees, J. (1972). The Optimum Town. *The Swedish Journal of Economics*, 114–135.

Miyauchi, Y., K. Nakajima, and S. J. Redding (2025). The economics of spatial mobility: Theory and evidence using smartphone data. *The Quarterly Journal of Economics 140*(4), 2507–2570.

Mongey, S. and M. Waugh (2024). Discrete Choice, Complete Markets, and Equilibrium. Working Paper 32135, National Bureau of Economic Research.

Redding, S. (2016). Goods Trade, Factor Mobility, and Welfare. *Journal of International Economics 101*, 148–167.

Redding, S. (2025). Evaluating Transport Improvements in Spatial Equilibrium. *Working Paper*.

Redding, S. and E. Rossi-Hansberg (2017). Quantitative Spatial Economics. *Annual Review of Economics 9*, 21–58.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *Journal of Political Economy 90*(6), 1257–1278.

Rosen, S. (1979). Wage-Based Indexes of Urban Quality of Life. *Current Issues in Urban Economics*, 74–104.

Rossi-Hansberg, E., P.-D. Sarte, and F. Schwartzman (2023). Cognitive Hubs and Spatial Redistribution. Working Paper 26267, National Bureau of Economic Research.

Ruggles, S., S. Flood, R. Goeken, J. Grover, E. Meyer, J. Pacas, and M. Sobek (2023). IPUMS USA: Version 14.0 [dataset]. *Minneapolis, MN: IPUMS 10*, D010.

Small, K. and E. Verhoef (2007). *The Economics of Urban Transportation.* Routledge.

Solow, R. (1957). Technical Change and the Aggregate Production Function. *The Review of Economics and Statistics 39*(3), 312–320.

Tsivanidis, N. (2025). Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota's Transmilenio. *Working Paper*.

Werning, I. (2007). Pareto Efficient Income Taxation. *Working Paper*.

Zárate, R. (2024). Spatial Misallocation, Informality, and Transit Improvements. *Working Paper*.

Online Appendix for
“Unpacking Aggregate Welfare in a Spatial Economy”

Eric Donald, Masao Fukui, Yuhei Miyauchi

January 7, 2026

A Details on Proofs

A.1 Proof of Lemma 1

Economy with Heterogeneous Preferences. Consider the problem of households of type θ deciding where to live. We index each individual by $\omega \in \mathcal{H}^\theta \equiv [0, \ell^\theta]$, and $\{\epsilon_j^\theta(\omega)\}_j$ denotes the preference draw of individual ω. Each individual solves the following problem:

$$\begin{aligned} v^\theta(\omega) = \max_{\{\mathbb{I}_j^\theta(\omega)\}_j} \sum_j \mathbb{I}_j^\theta(\omega) \left[u_j^\theta(C_j^\theta) + \epsilon_j^\theta(\omega)\right] \\ \text{s.t.} \quad \sum_j \mathbb{I}_j^\theta(\omega) = 1, \end{aligned} \tag{A.1}$$

where $\mathbb{I}_j^\theta(\omega) \in \{0,1\}$ is an indicator function for the location choice of individual ω. The fraction of individuals living in location j is given by

$$\mu_j^\theta = \frac{1}{\ell^\theta} \int_0^{\ell^\theta} \mathbb{I}_j^\theta(\omega) d\omega. \tag{A.2}$$

Economy with Representative Agent. Define the following function:

$$\begin{aligned} \psi^\theta(\{\mu_j^\theta\}_j) = -\max_{\{\mathbb{I}_j^\theta(\omega)\}_{\omega,j}} \frac{1}{\ell^\theta} \int_0^{\ell^\theta} \sum_j \epsilon_j^\theta(\omega) \mathbb{I}_j^\theta(\omega) d\omega \\ \text{s.t.} \quad \frac{1}{\ell^\theta} \int_0^{\ell^\theta} \mathbb{I}_j^\theta(\omega) d\omega = \mu_j^\theta \\ \sum_j \mathbb{I}_j^\theta(\omega) = 1. \end{aligned} \tag{A.3}$$

The representative agent solves

$$W^\theta = \max_{\{\mu_j^\theta\}_j : \sum_j \mu_j^\theta = 1} \sum_j \mu_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\mu_j^\theta\}) \tag{A.4}$$

The representative agent's first-order condition for μ_j^θ follows

$$u_j^\theta(C_j^\theta) - \frac{\partial \psi^\theta}{\partial \mu_j^\theta} = \delta^\theta, \tag{A.5}$$

where δ^θ is the Lagrange multiplier on the adding up constraint for $\{\mu_j^\theta\}_j$.

Equivalence Result. We formally restate the equivalence result of Lemma 1 as follows.

Lemma. *Suppose $\{\mathbb{I}_j^\theta(\omega)\}_j$ solves (A.1) for all ω. Then, $\{\mu_j^\theta\}_j$, given by (A.2), solves (A.4). Conversely, suppose $\{\mu_j^\theta\}_j$ solves (A.4). Then $\{\mathbb{I}_j^\theta(\omega)\}_{\omega,j}$, given by the solution to (A.3) associated with*

$\{\mu_j^\theta\}_j$, *solves* (A.1) *for almost all* ω. *Moreover, expected utility in the economy with heterogeneous preferences equals the utility of the representative agent:*

$$\frac{1}{\ell^\theta}\int_0^{\ell^\theta} v^\theta(\omega)d\omega = W^\theta$$

Proof. We prove the first part. Suppose to the contrary, there exists $\{\tilde{\mu}_j^\theta\}_j$ such that

$$\sum_j \tilde{\mu}_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\tilde{\mu}_j^\theta\}) > \sum_j \mu_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\mu_j^\theta\}). \tag{A.6}$$

Let $\{\tilde{\mathbb{I}}_j^\theta(\omega)\}_{\omega,j}$ denote the solution to (A.3) associated with $\{\tilde{\mu}_j^\theta\}_j$. Plugging into (A.6),

$$\frac{1}{\ell^\theta}\int_0^{\ell^\theta}\sum_j \tilde{\mathbb{I}}_j^\theta(\omega)\left[u_j^\theta(C_j^\theta)+\epsilon_j^\theta(\omega)\right]d\omega > \frac{1}{\ell^\theta}\int_0^{\ell^\theta}\sum_j \mathbb{I}_j^\theta(\omega)\left[u_j^\theta(C_j^\theta)+\epsilon_j^\theta(\omega)\right]d\omega, \tag{A.7}$$

where $\sum_j \tilde{\mathbb{I}}_j^\theta(\omega) = 1$ and $\sum_j \mathbb{I}_j^\theta(\omega) = 1$ for all ω. However, this is a contradiction because by our presumption, for any ω,

$$\sum_j \mathbb{I}_j^\theta(\omega)\left[u_j^\theta(C_j^\theta)+\epsilon_j^\theta(\omega)\right] \geq \sum_j \tilde{\mathbb{I}}_j^\theta(\omega)\left[u_j^\theta(C_j^\theta)+\epsilon_j^\theta(\omega)\right]$$

for all $\tilde{\mathbb{I}}_j^\theta(\omega)$, which would imply

$$\frac{1}{\ell^\theta}\int_0^{\ell^\theta}\sum_j \tilde{\mathbb{I}}_j^\theta(\omega)\left[u_j^\theta(C_j^\theta)+\epsilon_j^\theta(\omega)\right]d\omega \leq \frac{1}{\ell^\theta}\int_0^{\ell^\theta}\sum_j \mathbb{I}_j^\theta(\omega)\left[u_j^\theta(C_j^\theta)+\epsilon_j^\theta(\omega)\right]d\omega. \tag{A.8}$$

Now we prove the converse. Suppose to the contrary, there exists $\{\tilde{\mathbb{I}}_j^\theta(\omega)\}_j$ such that

$$\sum_j \tilde{\mathbb{I}}_j^\theta(\omega)\left[u_j^\theta(C_j^\theta)+\epsilon_j^\theta(\omega)\right] > \sum_j \mathbb{I}_j^\theta(\omega)\left[u_j^\theta(C_j^\theta)+\epsilon_j^\theta(\omega)\right] \tag{A.9}$$

and $\sum_j \tilde{\mathbb{I}}_j^\theta(\omega) = 1$ hold for all $\omega \in \Omega$, where $\Omega \subset [0, \ell^\theta]$ and $|\Omega| > 0$. Define

$$\tilde{\mu}_j^\theta = \frac{1}{\ell^\theta}\int_0^{\ell^\theta}\tilde{\mathbb{I}}_j^\theta(\omega)d\omega. \tag{A.10}$$

Then

$$\begin{aligned}\sum_j \mu_j^\theta u_j^\theta(C_j^\theta) - \psi(\{\mu_j^\theta\}) &= \frac{1}{\ell^\theta}\int_0^{\ell^\theta}\sum_j \mathbb{I}_j^\theta(\omega)\left[u_j^\theta(C_j^\theta)+\epsilon_j(\omega)\right]d\omega \\ &< \frac{1}{\ell^\theta}\int_0^{\ell^\theta}\sum_j \tilde{\mathbb{I}}_j^\theta(\omega)\left[u_j^\theta(C_j^\theta)+\epsilon_j^\theta(\omega)\right]d\omega\end{aligned}$$

$$\leq \sum_j \tilde{\mu}_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\tilde{\mu}_j^\theta\}).$$

This is a contradiction that $\{\mu_j^\theta\}_j$ is a solution to (A.4).

We need to show that the expected utility in the two economies coincides. This immediately follows given the above result. Let $\{\mathbb{I}_j^\theta(\omega)\}_{\omega,j}$ be the solution to (A.1) for all ω, and let $\{\mu_j^\theta\}_j$ denote the solution to (A.4). Then

$$\frac{1}{\ell^\theta}\int_0^{\ell^\theta} \sum_j \mathbb{I}_j^\theta(\omega)\left[u_j^\theta(C_j^\theta) + \epsilon_j^\theta(\omega)\right] d\omega = \sum_j \mu_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\mu_j^\theta\}). \tag{A.11}$$

□

A.2 Proof of Lemma 2

The constraints (22) and (23) immediately imply that at the solution of the pseudo-planning problem, we have $C_j^\theta = \check{C}_j^\theta$ and $\mu_j^\theta = \check{\mu}_j^\theta$, and therefore $l_j^\theta = \check{l}_j^\theta$ as well. Furthermore, (24) indicates that productivity net of agglomeration and congestion spillovers coincide with those in the equilibrium, i.e. $\mathcal{A}_{ij,k} = \check{\mathcal{A}}_{ij,k}$.

We now show that the remaining allocation of the pseudo-planning problem coincides with the decentralized equilibrium. In the decentralized equilibrium, quantities $\{\{\check{\mathbf{c}}_j^\theta\}, \{\check{\mathbf{x}}_{ij,k}, \check{\mathbf{l}}_{ij,k}, \check{h}_{ij,k}\}\}$ and prices $\{\{P_j^\theta, w_j^\theta\}, \{p_{ij,k}\}, r_j\}$ solve the resource constraints (13)-(16) as well as the following firms' optimality conditions:

$$P_j^\theta \frac{\partial \mathcal{C}_j^\theta}{\partial c_{ij,k}^\theta} = p_{ij,k}, \quad p_{ij,k}\mathcal{A}_{ij,k}\frac{\partial f_{ij,k}}{\partial l_{ij,k}^\theta} = w_i^\theta, \quad p_{ij,k}\mathcal{A}_{ij,k}\frac{\partial f_{ij,k}}{\partial h_{ij,k}} = r_i, \quad p_{ij,k}\mathcal{A}_{ij,k}\frac{\partial f_{ij,k}}{\partial x_{ij,k}^{l,m}} = p_{li,m}. \tag{A.12}$$

The first-order conditions of the pseudo-planning problem with respect to $c_{ij,k}$, $l_{ij,k}^\theta$, $h_{ij,k}$, and $x_{ij,k}^{l,m}$ are

$$P_j^{L,\theta} \frac{\partial \mathcal{C}_j^\theta}{\partial c_{ij,k}^\theta} = p_{ij,k}^L, \quad p_{ij,k}^L\mathcal{A}_{ij,k}\frac{\partial f_{ij,k}}{\partial l_{ij,k}^\theta} = w_i^{L,\theta}, \quad p_{ij,k}^L\mathcal{A}_{ij,k}\frac{\partial f_{ij,k}}{\partial h_{ij,k}} = r_i^L, \quad p_{ij,k}^L\mathcal{A}_{ij,k}\frac{\partial f_{ij,k}}{\partial x_{ij,k}^{l,m}} = p_{li,m}^L, \tag{A.13}$$

where $\{\{P_j^{L,\theta}, w_j^{L,\theta}\}, \{p_{ij,k}^L\}, r_j^L\}$ are Lagrange multipliers on constraints (13)-(16). Therefore, the decentralized equilibrium allocation satisfies the optimality conditions of the pseudo-planning problem. Moreover, equilibrium prices and Lagrange multipliers in the pseudo-planning problem coincide up to a multiplicative constant.

□

A.3 Proof of Proposition 1

The Lagrangian for the pseudo-planning problem is

$$\begin{aligned}
\mathcal{L} = \mathcal{W}&\left(\left\{\sum_j \hat{\mu}_j^\theta(\boldsymbol{C}^\theta)u_j^\theta(C_j^\theta) - \psi^\theta(\{\hat{\mu}_j^\theta(\boldsymbol{C}^\theta)\})\right\}\right) \\
&+ \sum_{i,j,k} p_{ij,k}^L \left[A_{ij,k}g_{ij,k}(\{\ell^\theta\hat{\mu}_j^\theta(\boldsymbol{C}^\theta)\}_\theta, y_{ij,k})f_{ij,k}(\mathbf{l}_{ij,k}, h_{ij,k}, \mathbf{x}_{ij,k}) - \left(\sum_\theta c_{ij,k}^\theta + \sum_{l,m} x_{jl,m}^{i,k}\right)\right] \\
&+ \sum_{j,\theta} P_j^{L,\theta}\left[\mathcal{C}_j^\theta(\mathbf{c}_j^\theta) - C_j^\theta\ell^\theta\hat{\mu}_j^\theta(\boldsymbol{C}^\theta)\right] \\
&+ \sum_{j,\theta} w_j^{L,\theta}\left[\ell^\theta\hat{\mu}_j^\theta(\boldsymbol{C}^\theta) - \sum_{i,k} l_{ji,k}^\theta\right] \\
&+ \sum_j r_j^L\left[\bar{h}_j - \sum_{i,k} h_{ji,k}\right] \\
&+ \sum_{j,\theta}\eta_j^\theta\left[C_j^\theta - \check{C}_j^\theta\right] \\
&+ \sum_{i,j,k}\phi_{i,j,k}\left[\mathcal{A}_{i,j,k} - \check{\mathcal{A}}_{i,j,k}\right],
\end{aligned}$$

where we have substituted constraints (21) and (22). Since one of the constraints $C_j^\theta = \check{C}_j^\theta$ is implied by the others from Walras' law, we normalize $\{\eta_j^\theta\}$ such that $\sum_{j,\theta}\frac{1}{u_j^{\theta\prime}(\check{C}_j^\theta)}\eta_j^\theta = 0$.[1] The first-order condition of the pseudo-planning problem with respect to C_j^θ is given by

$$\begin{aligned}
&\Lambda^\theta\ell^\theta\mu_j^\theta u_j^{\theta\prime}(C_j^\theta) - P_j^{L,\theta}\ell^\theta\mu_j^\theta \\
&+ \sum_l \frac{\partial\hat{\mu}_l^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta}\left[w_l^{L,\theta}\ell^\theta - P_l^{L,\theta}\ell^\theta + \sum_{j,k} p_{lj,k}^L\mathcal{A}_{lj,k}f_{lj,k}(\mathbf{l}_{lj,k}, h_{lj,k}, \mathbf{x}_{lj,k})\gamma_{lj,k}^\theta\ell^\theta\frac{1}{l_l^\theta}\right] + \eta_j^\theta = 0.
\end{aligned}$$

Here, we have used the fact that $\frac{\partial\hat{\mu}_j^\theta(\boldsymbol{C}^\theta)}{\partial C_j}u_j^\theta(C_j^\theta) - \frac{\partial\psi^\theta(\{\hat{\mu}_j^\theta(\boldsymbol{C}^\theta)\})}{\partial C_j^\theta} = 0$, because of the envelope theorem applied to the households' location choice (Lemma 1). Dividing both sides of the equation

[1] Given equilibrium $\{\check{C}_j^\theta\}$ for $j \neq \tilde{j}$ and $\theta \neq \tilde{\theta}$ for some $\tilde{j}, \tilde{\theta}$, Walras' Law implies for a specific value for equilibrium $C_{\tilde{j}}^{\tilde{\theta}}$ under resource constraints. Since we impose the resource constraints in the pseudo-planning problem, the constraint on $C_{\tilde{j}}^{\tilde{\theta}}$ is redundant. Following a similar logic, we can drop the linear combination for the constraints for C_j^θ: $\sum_{j,\theta}\frac{1}{u_j^{\theta\prime}(\check{C}_j^\theta)}C_j^\theta = \sum_{j,\theta}\frac{1}{u_j^{\theta\prime}(\check{C}_j^\theta)}\check{C}_j^\theta$ by setting the associated multiplier to 0, which gives the normalization of η_j^θ above. We choose this particular normalization so that the Lagrange multipliers exactly coincide with equilibrium prices (not only up to a multiplicative constant), which helps our exposition below.

above by $u_j^{\theta\prime}(C_j^\theta)$ and adding up across j and θ, we have

$$\begin{aligned}
&\sum_\theta \Lambda^\theta \ell^\theta - \sum_\theta \sum_j \frac{P_j^{L,\theta}}{u_j^{\theta\prime}(C_j^\theta)} l_j^\theta \\
&+ \sum_\theta \sum_l \left[\sum_j \frac{\partial \hat{\mu}_l^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta} \frac{1}{u_j^{\theta\prime}(C_j^\theta)} \right] \left[w_l^{L,\theta} \ell^\theta - P_l^{L,\theta} \ell^\theta + \sum_{j,k} p_{lj,k}^L \mathcal{A}_{lj,k} f_{lj,k}(\mathbf{l}_{lj,k}, h_{lj,k}, \mathbf{x}_{lj,k}) \gamma_{lj,k}^\theta \ell^\theta \frac{1}{l_l^\theta} \right] \\
&+ \sum_\theta \sum_j \frac{1}{u_j^{\theta\prime}(C_j^\theta)} \eta_j^\theta = 0. \qquad \text{(A.14)}
\end{aligned}$$

Notice here that, from the households' location problem (Lemma 1), location choice probability as a function of utility $\tilde{\mu}_l^\theta(\{u_l^\theta(C_l^\theta)\}_l)(\equiv \hat{\mu}_l^\theta(\boldsymbol{C}^\theta))$ is a homogeneous function of degree 0. Therefore,

$$\sum_j \frac{\partial \hat{\mu}_l^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta} \frac{1}{u_j^{\theta\prime}(C_j^\theta)} = \sum_j \frac{\partial \tilde{\mu}_l^\theta(\{u_l^\theta\}_l)}{\partial u_j^\theta} = 0. \qquad \text{(A.15)}$$

Together with the aforementioned normalization $\sum_\theta \sum_j \frac{1}{u_j^{\theta\prime}(C_j^\theta)} \eta_j^\theta = 0$, equation (A.14) becomes

$$\sum_\theta \sum_j \frac{P_j^{L,\theta}}{u_j^{\theta\prime}(C_j^\theta)} l_j^\theta = \sum_\theta \Lambda^\theta \ell^\theta. \qquad \text{(A.16)}$$

Since this corresponds to our normalization of SWF (28) with $P_j^{L,\theta} = P_j^\theta$, the Lagrange multipliers exactly coincide with equilibrium prices (not only up to a multiplicative constant as stated in the proof of Lemma 2), we have $P_j^{L,\theta} = P_j^\theta$, $p_{ij,k}^L = p_{ij,k}$, $w_i^{L,\theta} = w_i^\theta$, and $r_i^L = r_i$.

By applying the envelope theorem, we have

$$\begin{aligned}
\frac{dW}{d \ln A_{il,k}} =& \frac{d\mathcal{L}}{d \ln A_{il,k}} \\
=& p_{il,k} y_{il,k} - \sum_\theta \sum_j \eta_j^\theta \frac{d\check{C}_j^\theta}{d \ln A_{il,k}} - \sum_{j,m,s} \phi_{i,j,k} \frac{d\mathcal{A}_{jm,s}}{d \ln A_{il,k}} \\
=& p_{il,k} y_{il,k} + \sum_\theta \sum_j l_j^\theta [\Lambda^\theta u_j^{\theta\prime}(C_j^\theta) - P_j^\theta] \frac{d\check{C}_j^\theta}{d \ln A_{il,k}} \\
&+ \sum_\theta \sum_j \left[w_j^\theta - P_j^\theta C_j^\theta + \sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^\theta} \gamma_{jl,k}^\theta \right] \ell^\theta \frac{\partial \hat{\mu}_j^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta} \frac{d\check{C}_j^\theta}{d \ln A_{il,k}} \\
&+ \sum_{j,m,s} p_{jm,s} y_{jm,s} \gamma_{jm,s}^Y \frac{d \ln y_{jm,s}}{d \ln A_{il,k}}.
\end{aligned}$$

Note that, to obtain the last term, we used the fact that $\phi_{i,j,k} = -p_{ij,k} f_{ij,k}(\mathbf{l}_{ij,k}, h_{ij,k}, \mathbf{x}_{ij,k})$ from the first-order condition with respect to $\mathcal{A}_{jj,k}$. Therefore, by noting that $dl^\theta_j = \ell^\theta d\mu^\theta_j$, we have

$$dW = \sum_{i,j,k} p_{ij,k} y_{ij,k} d\ln A_{ij,k} + \sum_\theta \sum_j l^\theta_j [\Lambda^\theta u^{\theta\prime}_j(C^\theta_j) - P^\theta_j] dC^\theta_j$$
$$+ \sum_\theta \sum_j [w^\theta_j - P^\theta_j C^\theta_j] dl^\theta_j + \sum_{j,l,k} p_{jl,k} y_{jl,k} \left(\sum_\theta \gamma^\theta_{jl,k} d\ln l^\theta_j + \gamma^Y_{jl,k} d\ln y_{jl,k} \right). \quad \text{(A.17)}$$

Now,

$$\begin{aligned}
&\sum_\theta \sum_j l^\theta_j [\Lambda^\theta u^{\theta\prime}_j(C^\theta_j) - P^\theta_j] dC^\theta_j \\
=&\sum_\theta \ell^\theta \sum_j \mu^\theta_j [\Lambda^\theta - \frac{P^\theta_j}{u^{\theta\prime}_j(C^\theta_j)}] u^{\theta\prime}_j(C^\theta_j) dC^\theta_j \\
=&\sum_\theta \ell^\theta \left[\mathrm{Cov}_{j|\theta}(\Lambda^\theta - \frac{P^\theta_j}{u^{\theta\prime}_j(C^\theta_j)}, u^{\theta\prime}_j(C^\theta_j) dC^\theta_j) + \mathbb{E}_{j|\theta}[\Lambda^\theta - \frac{P^\theta_j}{u^{\theta\prime}_j(C^\theta_j)}] \mathbb{E}_{j|\theta}[u^{\theta\prime}_j(C^\theta_j) dC^\theta_j] \right] \\
=&\sum_\theta \ell^\theta \left[\mathrm{Cov}_{j|\theta}(-\frac{P^\theta_j}{u^{\theta\prime}_j(C^\theta_j)}, u^{\theta\prime}_j(C^\theta_j) dC^\theta_j) + \left(\Lambda^\theta - \mathbb{E}_{j|\theta}[\frac{P^\theta_j}{u^{\theta\prime}_j(C^\theta_j)}] \right) \mathbb{E}_{j|\theta}[u^{\theta\prime}_j(C^\theta_j) dC^\theta_j] \right] \\
=&\mathbb{E}_\theta[\mathrm{Cov}_{j|\theta}(-\frac{P^\theta_j}{u^{\theta\prime}_j(C^\theta_j)}, u^{\theta\prime}_j(C^\theta_j) dC^\theta_j)] + \mathrm{Cov}_\theta \left(\Lambda^\theta - \mathbb{E}_{j|\theta}[\frac{P^\theta_j}{u^{\theta\prime}_j(C^\theta_j)}], \mathbb{E}_{j|\theta}[u^{\theta\prime}_j(C^\theta_j) dC^\theta_j] \right) \\
&+ \mathbb{E}_\theta \left[\Lambda^\theta - \mathbb{E}_{j|\theta}[\frac{P^\theta_j}{u^{\theta\prime}_j(C^\theta_j)}] \right] \mathbb{E}_\theta \left[\mathbb{E}_{j|\theta}[u^{\theta\prime}_j(C^\theta_j) dC^\theta_j] \right] \\
=&\mathbb{E}_\theta[\mathrm{Cov}_{j|\theta}(-\frac{P^\theta_j}{u^{\theta\prime}_j(C^\theta_j)}, u^{\theta\prime}_j(C^\theta_j) dC^\theta_j)] + \mathrm{Cov}_\theta \left(\Lambda^\theta - \mathbb{E}_{j|\theta}[\frac{P^\theta_j}{u^{\theta\prime}_j(C^\theta_j)}], \mathbb{E}_{j|\theta}[u^{\theta\prime}_j(C^\theta_j) dC^\theta_j] \right),
\end{aligned}$$

where the third to last equation used the fact that $\mathrm{Cov}_{j|\theta}(X^\theta, Y^\theta_j) = 0$ for any variable X^θ, Y^θ_j where X^θ does not depend on j, and the last equation used the fact that $\mathbb{E}_\theta\left[\Lambda^\theta\right] = \mathbb{E}_\theta\left[\mathbb{E}_{j|\theta}[\frac{P^\theta_j}{u^{\theta\prime}_j(C^\theta_j)}]\right] = \mathbb{E}_\theta\left[\frac{P^\theta_j}{u^{\theta\prime}_j(C^\theta_j)}\right]$ under our price normalization (28). The two terms correspond to the (ii) MU dispersion and (v) redistribution terms in Proposition 1.

Similarly,

$$\begin{aligned}
&\sum_\theta \sum_j [w^\theta_j - P^\theta_j C^\theta_j] dl^\theta_j \\
=& \sum_\theta \sum_j \ell^\theta \mu^\theta_j [w^\theta_j - P^\theta_j C^\theta_j] d\ln l^\theta_j \\
=& \mathrm{Cov}_{j,\theta}\left(w^\theta_j - P^\theta_j C^\theta_j, d\ln l^\theta_j\right) + \mathbb{E}_{j,\theta}\left[w^\theta_j - P^\theta_j C^\theta_j\right] \underbrace{E_{j,\theta}\left[d\ln l^\theta_j\right]}_{=0} \\
=& \mathrm{Cov}_{j,\theta}\left(w^\theta_j - P^\theta_j C^\theta_j, d\ln l^\theta_j\right)
\end{aligned}$$

$$= \text{Cov}_{j,\theta}\left(-\Pi^\theta - T_j^\theta, d\ln l_j^\theta\right)$$
$$= \text{Cov}_{j,\theta}\left(-T_j^\theta, d\ln l_j^\theta\right)$$

which corresponds to the (iii) fiscal externality term. Finally,

$$\sum_\theta \sum_j \sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^\theta} \gamma_{jl,k}^\theta dl_j^\theta$$
$$= \sum_\theta \sum_j \ell^\theta \mu_j^\theta \sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^\theta} \gamma_{jl,k}^\theta d\ln l_j^\theta$$
$$= \text{Cov}_{j,\theta}\left(\sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^\theta} \gamma_{jl,k}^\theta, d\ln l_j^\theta\right).$$

This term plus $\sum_{j,l,k} p_{jl,k} y_{jl,k} \gamma_{jl,k}^Y d\ln y_{jl,k}$ corresponds to the (iv) technological externality term.

□

A.4 Proof of Proposition 2 and its Generalization

This Appendix provides the proof of Proposition 2 (Appendix A.4.1 and A.4.2). It also discusses a variant of Proposition 2 with degenerate idiosyncratic location preferences (Appendix A.4.3), as well as the case with output-based externalities and subsidies/taxes (i.e. $\gamma_{ij,k}^Y \neq 0$; Appendix A.4.4).

A.4.1 Heuristic Proof of Proposition 2

We first provide a heuristic proof using Proposition 1. A perturbation of dC_i^θ leads to

$$\Omega_{MU} = -\ell^\theta \mu_i^\theta P_i^\theta dC_i^\theta + \mathbb{E}_{j|\theta}\left[P_j^\theta / u_j^{\theta\prime}(C_j^\theta)\right] \ell^\theta \mu_i^\theta u_i^{\theta\prime}(C_i^\theta) dC_i^\theta,$$

and

$$\Omega_R = \ell^\theta \mu_i^\theta \Lambda^\theta u_i^{\theta\prime}(C_i^\theta) dC_i^\theta - \mathbb{E}_{j|\theta}\left[P_j^\theta / u_j^{\theta\prime}(C_j^\theta)\right] \ell^\theta \mu_i^\theta u_i^{\theta\prime}(C_i^\theta) dC_i^\theta.$$

The optimality of the spatial transfers requires $dW = \Omega_{MU} + \Omega_{FE} + \Omega_{TE} + \Omega_R = 0$. By noting that $d\ln l_j^\theta = \frac{\partial \ln \mu_j^\theta(\boldsymbol{C}^\theta)}{\partial C_i^\theta} dC_i^\theta$, we obtain equation (31).

A.4.2 Formal Proof of Proposition 2

Consider the government's problem to set spatial transfers T_j^θ to maximize the SWF (17). The (constrained) Pareto efficient transfer policy solves

$$\max_{\{\{C_j^\theta, \mathbf{c}_j^\theta, P_j^\theta, T_j^\theta, l_j^\theta, \mu_j^\theta, w_j^\theta\}, W^\theta\}, \{\mathbf{x}_{ij,k}, \mathbf{l}_{ij,k}, h_{ij,k}, \mathcal{A}_{ij,k}, p_{ij,k}\}, r_j\}} \sum_j \mathcal{W}(\{W^\theta\}), \tag{A.18}$$

subject to (2)-(16).

To solve this problem, we follow the primal approach in the public finance literature. That is, we focus on a relaxed planning problem where the Planner picks an incentive-compatible allocation and later confirms that their chosen allocation, alongside supporting prices, is also a solution to the original problem. The relaxed planning problem is defined as follows.

Definition A.1 (Relaxed Planning Problem)**.** The Planner solves

$$\max_{\{\{C_j^\theta, \mathbf{c}_j^\theta, l_j^\theta, \mu_j^\theta\}, \{\mathbf{x}_{ij,k}, \mathbf{l}_{ij,k}, h_{ij,k}, \mathcal{A}_{ij,k}\}\}} \mathcal{W}(\{W^\theta\}) \tag{A.19}$$

$$\text{subject to (9), (13)-(16)}$$

$$\{\mu_j^\theta\}_j \in \underset{\{\tilde{\mu}_j^\theta\}_j : \sum_j \tilde{\mu}_j^\theta = 1}{\arg\max} \sum_j \tilde{\mu}_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\tilde{\mu}_j^\theta\}) \tag{A.20}$$

Compared to the pseudo-planning problem in Lemma 2, the relaxed planning problem chooses consumption instead of taking the equilibrium allocation as given (equation (23)). Furthermore, this problem is different from the first-best planning problem as considered in Appendix B because the Planner must choose an incentive-compatible population allocation (A.20). For this reason, we refer to these policies as second-best.

We first characterize the first-order conditions of the relaxed planning problem of Definition A.1. The first-order conditions with respect to $c_{ij,k}^\theta$, $l_{ij,k}^\theta$, $h_{ij,k}$, and $x_{ij,k}^{l,m}$ are given by

$$P_j^{SB,\theta} \frac{\partial \mathcal{C}_j^\theta}{\partial c_{ij,k}^\theta} = p_{ij,k}^{SB}, \quad p_{ij,k}^{SB} \mathcal{A}_{ij,k} \frac{\partial f_{ij,k}}{\partial l_{ij,k}^\theta} = w_i^{SB,\theta}, \quad p_{ij,k}^{SB} \mathcal{A}_{ij,k} \frac{\partial f_{ij,k}}{\partial h_{ij,k}} = r_i^{SB}, \quad p_{ij,k}^{SB} \mathcal{A}_{ij,k} \frac{\partial f_{ij,k}}{\partial x_{ij,k}^{l,m}} = p_{li,m}^{SB}, \tag{A.21}$$

where $P_j^{SB,\theta}$, $p_{ij,k}^{SB}$, $w_j^{SB,\theta}$, and r_j^{SB} are Lagrange multipliers on constraints (13)-(16). These conditions are identical to the equilibrium conditions (A.12), with $P_j^{SB,\theta}$, $p_{ij,k}^{SB}$, $w_j^{SB,\theta}$, and r_j^{SB} coinciding with P_j^θ, $p_{ij,k}$, w_j^θ, and r_j up to a multiplicative constant. The first-order condition with respect to C_j^θ is given by

$$\ell^\theta \mu_j^\theta \left[\Lambda^\theta u_j^{\theta\prime}(C_j^\theta) - P_j^{SB,\theta}\right] = \ell^\theta \sum_i \frac{\partial \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta} \left[P_i^{SB,\theta} C_i^\theta - w_i^{SB,\theta} - \sum_{l,k} p_{il,k}^{SB} y_{il,k} \frac{1}{l_i^\theta} \gamma_{il,k}^\theta\right]. \tag{A.22}$$

Dividing both sides by $u_j^{\theta\prime}(C_j^\theta)$ and summing across j and θ, we have

$$\begin{aligned}
&\sum_\theta \ell^\theta \sum_j \mu_j^\theta \left[\Lambda^\theta - \frac{P_j^{SB,\theta}}{u_j^{\theta\prime}(C_j^\theta)}\right] \\
&= -\sum_\theta \ell^\theta \sum_j \sum_i \frac{1}{u_j^{\theta\prime}(C_j^\theta)} \frac{\partial \mu_i^\theta}{\partial C_j^\theta} \left[w_i^{SB,\theta} - P_i^{SB,\theta} C_i^\theta + \sum_{l,k} p_{il,k}^{SB} y_{il,k} \frac{1}{l_i^\theta} \gamma_{il,k}^\theta\right] \\
&= -\sum_\theta \ell^\theta \sum_i \underbrace{\sum_j \frac{1}{u_j^{\theta\prime}(C_j^\theta)} \frac{\partial \mu_i^\theta}{\partial C_j^\theta}}_{=0} \left[w_i^{SB,\theta} - P_i^{SB,\theta} C_i^\theta + \sum_{l,k} p_{il,k}^{SB} y_{il,k} \frac{1}{l_i^\theta} \gamma_{il,k}^\theta\right] \\
&= 0,
\end{aligned}$$

where the third line comes from the fact that a uniform increase in utility in all locations will not affect location choices. This implies

$$\sum_{\theta}\sum_{j} l_j^{\theta}\frac{P_j^{SB,\theta}}{u_j^{\theta\prime}(C_j^{\theta})}=\sum_{\theta}\ell^{\theta}\Lambda^{\theta}. \tag{A.23}$$

Recall that we have already established $P_j^{SB,\theta}$, $p_{ij,k}^{SB}$, $w_j^{SB,\theta}$, and r_j^{SB} coincide with P_j^{θ}, $p_{ij,k}$, w_j^{θ}, and r_j up to a multiplicative constant. Comparing (A.23) and (28) shows that this multiplicative constant is also one under the chosen numeraire. Therefore, we can rewrite (A.22) using equilibrium prices as

$$\mu_j^{\theta}\left[\Lambda^{\theta}u_j^{\theta\prime}(C_j^{\theta})-P_j^{\theta}\right]=\sum_i\frac{\partial\hat{\mu}_i^{\theta}(\boldsymbol{C}^{\theta})}{\partial C_j^{\theta}}\left[P_i^{\theta}C_i^{\theta}-w_i^{\theta}-\sum_{l,k}p_{il,k}y_{il,k}\frac{1}{l_i^{\theta}}\gamma_{il,k}^{\theta}\right]. \tag{A.24}$$

By noting $T_i^{\theta}=P_i^{\theta}C_i^{\theta}-w_i^{\theta}-\Pi^{\theta}$ and $\sum_i\frac{\partial\hat{\mu}_i^{\theta}}{\partial C_j^{\theta}}=0$, we obtain (31).

Next, we need to show that all the equilibrium conditions are satisfied under $T_j^{\theta}=P_j^{\theta}C_j^{\theta}-w_j^{\theta}-\Pi^{\theta}$ where C_j^{θ} satisfies (31) with supporting equilibrium prices $\{\{P_j^{\theta},w_j^{\theta}\},\{p_{ij,k}\},r_j\}$. First, it is immediate to see that market clearing conditions are satisfied because (13)-(16) enter as constraints. The constraint (A.20) implies that the population distribution solves (19). Given prices $\{\{P_j^{\theta},w_j^{\theta}\},\{p_{ij,k}\},r_j\}$, the firm's optimality conditions (A.13) are satisfied because they are identical to (A.21).

Finally, it remains to be shown that the government budget (8) is satisfied. Multiplying $T_j^{\theta}=P_j^{\theta}C_j^{\theta}-w_j^{\theta}-\Pi^{\theta}$ by l_j^{θ} and summing across j and θ, we have

$$\begin{aligned}
&\sum_{\theta}\sum_{j}T_j^{\theta}l_j^{\theta}\\
&=\sum_{\theta}\sum_{j}P_j^{\theta}C_j^{\theta}l_j^{\theta}-\sum_{\theta}\sum_{j}w_j^{\theta}l_j^{\theta}-\sum_{\theta}\ell^{\theta}\Pi^{\theta}\\
&=\sum_{\theta}\sum_{i,j,k}p_{ij,k}c_{ij,k}^{\theta}l_j^{\theta}-\sum_{\theta}\sum_{j}w_j^{\theta}l_j^{\theta}-\sum_{\theta}\ell^{\theta}\Pi^{\theta}\\
&=\sum_{i,j,k}p_{ij,k}\left[\mathcal{A}_{ij,k}f_{ij,k}(\mathbf{l}_{ij,k},h_{ij,k},\mathbf{x}_{ij,k})-\sum_{l,m}p_{ij,k}x_{jl,m}^{i,k}\right]\\
&\qquad-\sum_{\theta}\sum_{j}w_j^{\theta}l_j^{\theta}-\sum_{\theta}\ell^{\theta}\Pi^{\theta}\\
&=\sum_{i,j,k}p_{ij,k}\left[\sum_{\theta}\mathcal{A}_{ij,k}\frac{\partial f_{ij,k}}{\partial l_{ij,k}^{\theta}}l_{ij,k}^{\theta}+\mathcal{A}_{ij,k}\frac{\partial f_{ij,k}}{\partial h_{ij,k}}h_{ij,k}+\sum_{l,m}\mathcal{A}_{ij,k}\frac{\partial f_{ij,k}}{\partial x_{ij,k}^{l,m}}x_{ij,k}^{l,m}-\sum_{l,m}p_{ij,k}x_{jl,m}^{i,k}\right]\\
&\qquad-\sum_{\theta}\sum_{j}w_j^{\theta}l_j^{\theta}-\sum_{\theta}\ell^{\theta}\Pi^{\theta}\\
&=\sum_{i,j,k}\left[\sum_{\theta}w_i^{\theta}l_{ij,k}^{\theta}+r_ih_{ij,k}+\sum_{l,m}p_{li,m}x_{ij,k}^{l,m}-\sum_{l,m}p_{ij,k}x_{jl,m}^{i,k}\right]
\end{aligned}$$

$$
\begin{aligned}
&- \sum_\theta \sum_j w_j^\theta l_j^\theta - \sum_\theta \ell^\theta \Pi^\theta \\
&= \sum_\theta \sum_j w_j^\theta l_j + \sum_j r_j \bar{h}_j - \sum_\theta \sum_j w_j^\theta l_j^\theta - \sum_\theta \ell^\theta \Pi^\theta \\
&= 0.
\end{aligned}
$$

□

A.4.3 Proof of Proposition 2 With Degenerate Idiosyncratic Location Preferences

We also discuss Proposition 2 for the case with degenerate idiosyncratic location preferences as considered by Fajgelbaum and Gaubert (2020). To do so, we rewrite the second-best problem as follows:

$$
\max_{\{W^\theta,\{C_j^\theta,\mathbf{c}_j^\theta,\mu_j^\theta\}\},\{\mathbf{x}_{ij,k},\mathbf{l}_{ij,k},h_{ij,k},\mathcal{A}_{ij,k}\}} \sum_j \mu_j^{\bar{\theta}} u_j(C_j^{\bar{\theta}}) \tag{A.25}
$$

$$
\text{s.t.} \qquad (9), (13)\text{-}(16) \tag{A.26}
$$

$$
u_j(C_j^\theta) = W^\theta \text{ for all } j, \theta \tag{A.27}
$$

$$
\sum_j \mu_j^\theta = 1 \text{ for all } \theta \tag{A.28}
$$

$$
\sum_j \mu_j^{\tilde{\theta}} u_j(C_j^{\tilde{\theta}}) \geq \underline{W}^{\tilde{\theta}} \text{ for all } \tilde{\theta} \neq \bar{\theta} \tag{A.29}
$$

Note that we rewrote households' incentive compatibility constraints for location choice (22) with utility equalization (A.27) and adding up constraint (A.28). Note also that $\psi^\theta(\cdot) = 0$ with degenerate idiosyncratic location preferences.

The first-order condition for μ_j^θ is given by

$$
\ell^\theta \left[w_j^\theta - P_j^\theta C_j^\theta + \sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^\theta} \gamma_{jl,k}^\theta \right] + \hat{\Lambda}^\theta W^\theta = \delta^{SB,\theta}
$$

$$
\Leftrightarrow \quad w_j^\theta - P_j^\theta C_j^\theta + \sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^\theta} \gamma_{jl,k}^\theta = \left(\delta^{SB,\theta} - \hat{\Lambda}^\theta W^\theta \right) / \ell^\theta, \tag{A.30}
$$

where $\delta^{SB,\theta}$ denotes the Lagrange multiplier on constraint (A.28). By noting that $T_j^\theta = P_j^\theta C_j^\theta - w_j^\theta - \Pi^\theta$, the cross-location component of transfers only addresses technological externalities, and the cross-type component of transfers addresses redistribution concerns, as highlighted by Fajgelbaum and Gaubert (2020).

A.4.4 Generalization with Output-Based Externalities and Taxes/Subsidies

For expositional purposes, in Proposition 2 of the paper, we provide the optimal transfer formula under the assumption that the agglomeration externality depends only on local population size but not on output, i.e. $\gamma_{ij,k}^Y = 0$. In this section, we relax this assumption, but instead assume that the Planner can implement the output taxes/subsidies. We show that our formula of Proposition 2 remains unchanged if the Planner simultaneously implements the Pigouvian output tax/subsidies.

To consider such an environment, we generalize the environment as follows. We generalize the intermediate good producer's problem from (12) of our paper to

$$(\mathbf{l}_{ij,k}, h_{ij,k}, \mathbf{x}_{ij,k}) \in \underset{\tilde{\mathbf{l}}_{ij,k}, \tilde{h}_{ij,k}, \tilde{\mathbf{x}}_{ij,k}}{\arg\max} \; (1+\xi_{ij,k}) p_{ij,k} \mathcal{A}_{ij,k} f_{ij,k}(\tilde{\mathbf{l}}_{ij,k}, \tilde{h}_{ij,k}, \tilde{\mathbf{x}}_{ij,k}) - \sum_{\theta} w_i^{\theta} \tilde{l}_{ij,k}^{\theta} - r_i \tilde{h}_{ij,k} - \sum_{l,m} p_{li,m} \tilde{x}_{ij,k}^{l,m}, \tag{A.31}$$

where $\xi_{ij,k}$ is the rate of output subsidies set by the national government.

We also generalize the government budget constraint from (8) such that

$$\sum_{\theta} \sum_{j} T_j^{\theta} l_j^{\theta} + \sum_{i,j,k} \xi_{ij,k} p_{ij,k} y_{ij,k} = 0. \tag{A.32}$$

The decentralized equilibrium of this economy is defined by the prices $\{\{P_j^{\theta}, w_j^{\theta}\}, \{p_{ij,k}\}, r_j\}$, quantities $\{\{C_j^{\theta}, \mathbf{c}_j^{\theta}, \mu_j^{\theta}, l_j^{\theta}\}, \{\mathbf{x}_{ij,k}, \mathbf{l}_{ij,k}\}\}$, transfers $\{T_j^{\theta}\}$, output subsidies $\{\xi_{ij,k}\}$, and productivities $\{\mathcal{A}_{ij,k}\}$ that satisfy equations (2)-(7), (9)-(11), (13)-(16), (A.31), and (A.32).

Proposition A.1. *Consider an economy with output subsidies as defined above. Assume also that idiosyncratic location preferences are non-degenerate. The optimal spatial transfers must satisfy*

$$\mu_i^{\theta} \left[\Lambda^{\theta} u_i^{\theta\prime}(C_i^{\theta}) - P_i^{\theta} \right] = Cov_{j|\theta} \left(T_j^{\theta} - \frac{1}{l_j^{\theta}} \sum_{l,k} p_{jl,k} y_{jl,k} \gamma_{jl,k}^{\theta}, \frac{\partial \ln \mu_j^{\theta}(\boldsymbol{C}^{\theta})}{\partial C_i^{\theta}} \right) \quad \textit{for all } i, \theta, \tag{A.33}$$

and the optimal output subsidy must satisfy

$$1 + \xi_{ij,k} = \left(1 - \gamma_{ij,k}^{Y}\right)^{-1} \quad \textit{for all } i, j, k. \tag{A.34}$$

Notice that equation (A.33) is equivalent to the expression in equation (2) of our main paper. Equation (A.34) is a standard Pigouvian output subsidy to correct for the agglomeration externality in output. Therefore, if the Pigouvian subsidies are simultaneously implemented, our optimal transfer formula remains unchanged from the case without output externalities, i.e. $\gamma_{ij,k}^{Y} = 0$.

Proof. We follow the same proof structure for the original proposition, as discussed in Appendix A.4.2. That is, we focus on a relaxed planning problem where the Planner picks an incentive-compatible allocation and later confirms that their chosen allocation, alongside supporting prices, is also a solution to the original problem. For expositional purposes, we slightly rewrite the relaxed planning problem from Definition A.1:

$$\max_{\{\{C_j^{\theta}, \mathbf{c}_j^{\theta}, l_j^{\theta}, \mu_j^{\theta}\}, \{y_{ij,k}, \mathbf{x}_{ij,k}, \mathbf{l}_{ij,k}, h_{ij,k}, \mathcal{A}_{ij,k}\}\}} \mathcal{W}(\{W^{\theta}\}) \tag{A.35}$$

subject to (9), (14)-(16)

$$\sum_{\theta} c^{\theta}_{ij,k} + \sum_{l,m} x^{i,k}_{jl,m} = y_{ij,k} \tag{A.36}$$

$$y_{ij,k} = \mathcal{A}_{ij,k} f_{ij,k}(\mathbf{l}_{ij,k}, h_{ij,k}, \mathbf{x}_{ij,k}) \tag{A.37}$$

$$\{\mu^{\theta}_j\}_j \in \underset{\{\tilde{\mu}^{\theta}_j\}_j : \sum_j \tilde{\mu}^{\theta}_j = 1}{\arg\max} \sum_j \tilde{\mu}^{\theta}_j u^{\theta}_j(C^{\theta}_j) - \psi^{\theta}(\{\tilde{\mu}^{\theta}_j\}) \tag{A.38}$$

The first-order conditions of the relaxed planning problem of the relaxed problem with respect to $c^{\theta}_{ij,k}$, $l^{\theta}_{ij,k}$, $h_{ij,k}$, and $x^{l,m}_{ij,k}$ are given by

$$P^{SB,\theta}_j \frac{\partial \mathcal{C}^{\theta}_j}{\partial c^{\theta}_{ij,k}} = p^{SB}_{ij,k}, \quad \tilde{p}^{SB}_{ij,k} \mathcal{A}_{ij,k} \frac{\partial f_{ij,k}}{\partial l^{\theta}_{ij,k}} = w^{SB,\theta}_i, \quad \tilde{p}^{SB}_{ij,k} \mathcal{A}_{ij,k} \frac{\partial f_{ij,k}}{\partial h_{ij,k}} = r^{SB}_i, \quad \tilde{p}^{SB}_{ij,k} \mathcal{A}_{ij,k} \frac{\partial f_{ij,k}}{\partial x^{l,m}_{ij,k}} = p^{SB}_{li,m}, \tag{A.39}$$

where $p^{SB}_{ij,k}$, $\tilde{p}^{SB}_{ij,k}$, $P^{SB,\theta}_j$, $w^{SB,\theta}_j$, and r^{SB}_j are Lagrange multipliers on constraints (A.36), (A.37), (14)-(16). Furthermore, the first-order conditions with respect to $y_{ij,k}$ is given by

$$p^{SB}_{ij,k} = \tilde{p}^{SB}_{ij,k} - \tilde{p}^{SB}_{ij,k} \frac{\partial \mathcal{A}_{ij,k}}{\partial y_{ij,k}} f_{ij,k}(\mathbf{l}_{ij,k}, h_{ij,k}, \mathbf{x}_{ij,k}) \iff p^{SB}_{ij,k} = \tilde{p}^{SB}_{ij,k} \left[1 - \gamma^{Y}_{ij,k}\right] \tag{A.40}$$

The remaining conditions are unchanged from the main proof (Appendix A.4.2).

It is straightforward to follow the same step as in Appendix A.4.2 to show that the optimal transfers T^{θ}_j satisfy equation (A.33). It is also straightforward to see that the Pigouvian subsidies $1 + \xi_{ij,k} = \left(1 - \gamma^{Y}_{ij,k}\right)^{-1}$, as given by equation (A.34), can implement the equations (A.39) and (A.40). □

A.5 Proof of Proposition 3

By multiplying equation (31) by $\ell^{\theta} dC^{\theta}_j$ and summing up across j and θ, we have

$$\begin{aligned}
&\sum_j \sum_{\theta} l^{\theta}_j \left[\Lambda^{\theta} u^{\theta\prime}_j(C^{\theta}_j) - P^{\theta}_j\right] dC^{\theta}_j \\
&= \sum_i \sum_{\theta} \underbrace{\sum_j \ell^{\theta} dC^{\theta}_j \frac{\partial \mu^{\theta}_i}{\partial C^{\theta}_j}}_{= dl^{\theta}_i} \left[P^{\theta}_i C^{\theta}_i - w^{\theta}_i - \sum_{l,k} p_{il,k} y_{il,k} \frac{1}{l^{\theta}_i} \gamma^{\theta}_{il,k}\right].
\end{aligned} \tag{A.41}$$

Plugging this into equation (A.17), we only have (i) technology term.

One can also prove Proposition 3 by directly applying the envelope theorem to the relaxed planning problem of Definition A.1. Despite the presence of incentive compatibility constraints for households' location decisions, there are no reallocation effects because technological effects do not directly affect these constraints.

□

A.6 Proof of Proposition 4

From Baqaee and Farhi (2024),

$$d \ln Y_i = \sum_{j,k} \frac{p_{ij,k} y_{ij,k}}{GDP_i} d \ln \mathcal{A}_{ij,k} + \sum_{\theta} \frac{w_i^\theta l_i^\theta}{GDP_i} d \ln l_i^\theta. \tag{A.42}$$

Using the definition of aggregate real GDP such that $d \ln Y = \sum_i GDP_i d \ln Y_i$, and using the fact that $d \ln \mathcal{A}_{ij,k} = d \ln A_{ij,k} + \sum_\theta \gamma_{ij,k}^\theta d \ln l_i^\theta + \gamma_{ij,k}^Y d \ln y_{ij,k}$,

$$\begin{aligned} d \ln Y = & \sum_{i,j,k} p_{ij,k} y_{ij,k} d \ln A_{ij,k} + \sum_{i,j,k} \sum_{\theta} \gamma_{ij,k}^\theta p_{ij,k} y_{ij,k} d \ln l_i^\theta \\ & + \sum_{i,j,k} p_{ij,k} y_{ij,k} \gamma_{ij,k}^Y d \ln y_{ij,k} + \sum_{i} \sum_{\theta} w_i^\theta l_i^\theta d \ln l_i^\theta, \end{aligned} \tag{A.43}$$

which gives the desired expression.

□

A.7 Proof of Proposition 5

Here, we prove a general version of Proposition 5 nesting multiple household types.

Proposition A.2. *Consider the spatial equilibrium with multiplicative Fréchet preference shocks with arbitrary correlation (37) with multiple types. Let $\tilde{W} \equiv \tilde{\mathcal{W}}(\{\tilde{W}^\theta\})$ be welfare in this economy, where $\tilde{W}^\theta = \int_{\omega \in \mathcal{H}^\theta} \tilde{u}_j^\theta(C_j^\theta) \tilde{\epsilon}_j^\theta(\omega) \mathbb{I}\left[j = j^\theta(\omega)\right] d\omega$, and $\tilde{\epsilon}_j^\theta(\omega)$ indicates household ω's idiosyncratic location preference for location j. Consider another economy where preferences are a log transformation of the first specification, i.e. the additively separable counterpart (38) with multiple types, and the remaining equilibrium conditions of Definition 1 are unchanged. Let $W \equiv \mathcal{W}(\{W^\theta\})$ be welfare in this economy, where $\mathcal{W}(\{W^\theta\}) \equiv \ln \tilde{\mathcal{W}}(\{\exp(W^\theta)\})$ is the social welfare function. Then, the welfare decomposition of Proposition 1 is identical in both economies up to a multiplicative constant.*

Before starting our proof, notice that Proposition A.2 comes down to Proposition 5 with a single household type.

We first note that, given the assumption that $\mathcal{W}(\{W^\theta\}) \equiv \ln \tilde{\mathcal{W}}(\{\exp(W^\theta)\})$, we have

$$dW = \sum_{\theta} \frac{\partial \mathcal{W}}{\partial W^\theta} dW^\theta, \qquad d \ln \tilde{W} = \sum_{\theta} \frac{\partial \mathcal{W}}{\partial W^\theta} d \ln \tilde{W}^\theta. \tag{A.44}$$

Therefore, to show $dW = d \ln \tilde{W}$, it is sufficient to show that the same isomorphism holds for the expected utility for each type θ, i.e. $dW^\theta = d \ln \tilde{W}^\theta$. The expected utility of households in (37) is given by

$$\tilde{W}^\theta = G^\theta(\tilde{u}_1^\theta(C_1^\theta)^{\nu^\theta}, \ldots, \tilde{u}_N^\theta(C_N^\theta)^{\nu^\theta})^{1/\nu^\theta}, \tag{A.45}$$

and that in (38) is given by

$$W^\theta = \frac{1}{\nu^\theta} \ln G^\theta(\exp(\nu^\theta u_1^\theta(C_1^\theta)), \ldots, \exp(\nu^\theta u_N^\theta(C_N^\theta))). \tag{A.46}$$

See Appendix C for a detailed mathematical derivation. Therefore, under $u_j^\theta(C_j^\theta) = \ln(\tilde{u}_j^\theta(C_j^\theta))$ and $\epsilon_j^\theta = \ln(\tilde{\epsilon}_j^\theta)$, we have $W^\theta = \ln \tilde{W}^\theta$.

Finally, we prove that the decomposition is also identical. The Lagrangian for the pseudo-planning problem in an economy with multiplicative preference shocks (37) is

$$
\begin{aligned}
\mathcal{L} = \tilde{\mathcal{W}} & \left(\left\{G^\theta(\tilde{u}_1^\theta(C_1^\theta)^{\nu^\theta}, \ldots, \tilde{u}_N^\theta(C_N^\theta)^{\nu^\theta})^{1/\nu^\theta}\right\}\right) \\
& + \sum_{i,j,k} p_{ij,k}^L \left[A_{ij,k} g_{ij,k}(\{\ell^\theta \hat{\mu}_j^\theta(\boldsymbol{C}^\theta)\}_\theta, y_{ij,k}) f_{ij,k}(\mathbf{l}_{ij,k}, h_{ij,k}, \mathbf{x}_{ij,k}) - \left(\sum_\theta c_{ij,k}^\theta + \sum_{l,m} x_{jl,m}^{i,k}\right)\right] \\
& + \sum_{j,\theta} P_j^{L,\theta} \left[\mathcal{C}_j^\theta(\mathbf{c}_j^\theta) - C_j^\theta \ell^\theta \mu_j^\theta(\boldsymbol{C}^\theta)\right] \\
& + \sum_{j,\theta} w_j^{L,\theta} \left[\ell^\theta \hat{\mu}_j^\theta(\boldsymbol{C}^\theta) - \sum_{i,k} l_{ji,k}^\theta\right] \\
& + \sum_j r_j^L \left[\bar{h}_j - \sum_{i,k} h_{ji,k}\right] \\
& + \sum_j \eta_j^\theta \left[C_j^\theta - \check{C}_j^\theta\right] \\
& + \sum_{i,j,k} \phi_{i,j,k} \left[\mathcal{A}_{i,j,k} - \check{\mathcal{A}}_{i,j,k}\right],
\end{aligned}
$$

where we normalize $\{\eta_j^\theta\}$ such that $\sum_\theta \sum_j \eta_j^\theta / u^{\theta\prime}(\check{C}_j^\theta) = 0$ as in the proof of Proposition 1. As in Proposition 1, the Lagrange multipliers $\{\{P_j^{L,\theta}, w_j^{L,\theta}\}, \{p_{ij,k}^L\}, r_j^L\}$ coincide with equilibrium prices up to a multiplicative constant. The first-order condition with respect to C_j^θ is

$$
\ell^\theta \mu_j^\theta \left[\frac{\partial \tilde{\mathcal{W}}}{\partial \tilde{W}^\theta} \tilde{W}^\theta \underbrace{\frac{\tilde{u}_j^{\theta\prime}(C_j^\theta)}{\tilde{u}_j^\theta(C_j^\theta)}}_{= u_j^{\theta\prime}(C_j^\theta)} - P_j^{L,\theta}\right] + \eta_j^\theta = \ell^\theta \sum_i \frac{\partial \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta} \left[P_i^{L,\theta} C_i^\theta - w_i^{L,\theta} - \sum_{l,k} p_{il,k}^L y_{il,k} \frac{1}{l_i^\theta} \gamma_{il,k}^\theta\right],
$$

where we used the fact that (see also Appendix C)

$$
\frac{\partial \tilde{W}^\theta}{\partial C_j^\theta} = \tilde{W}^\theta \frac{G_j^\theta(\tilde{u}_1^\theta(C_1^\theta)^{\nu^\theta}, \ldots, \tilde{u}_N^\theta(C_N^\theta)^{\nu^\theta})}{G^\theta(\tilde{u}_1(C_1^\theta)^{\nu^\theta}, \ldots, \tilde{u}_N^\theta(C_N^\theta)^{\nu^\theta})} \tilde{u}_j^{\theta\prime}(C_j^\theta) = \tilde{W}^\theta \mu_j^\theta \frac{\tilde{u}_j^{\theta\prime}(C_j^\theta)}{\tilde{u}_j^\theta(C_j^\theta)} = \tilde{W}^\theta \mu_j^\theta u_j^{\theta\prime}(C_j^\theta). \quad \text{(A.47)}
$$

Since $\Lambda^\theta = \frac{1}{\tilde{\mathcal{W}}} \frac{\partial \tilde{\mathcal{W}}}{\partial \tilde{W}^\theta} \tilde{W}^\theta$ under our assumption, we can rewrite the above expression as

$$
\ell^\theta \mu_j^\theta \left[\Lambda^\theta \tilde{\mathcal{W}} u_j^{\theta\prime}(C_j^\theta) - P_j^{L,\theta}\right] + \eta_j^\theta = \ell^\theta \sum_i \frac{\partial \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta} \left[P_i^{L,\theta} C_i^\theta - w_i^{L,\theta} - \sum_{l,k} p_{il,k}^L y_{il,k} \frac{1}{l_i^\theta} \gamma_{il,k}^\theta\right].
$$

Dividing both sides by $u_j^{\theta\prime}(C_j^\theta)$ and summing up across j and θ,

$$\sum_\theta \sum_j l_j^\theta \frac{P_j^{L,\theta}}{u_j^{\theta\prime}(C_j^\theta)} = \tilde{\mathcal{W}} \mathbb{E}_\theta \left[\Lambda^\theta\right]. \tag{A.48}$$

Comparing (28) and (A.48), the Lagrange multipliers coincide with equilibrium prices up to the multiplicative constant $\tilde{\mathcal{W}}$.

Applying the envelope theorem as in the proof of Proposition 1,

$$\begin{aligned} d\tilde{W} =& \sum_{i,j,k} p_{ij,k}^L y_{ij,k} d\ln A_{ij,k} + \sum_\theta \sum_j l_j^\theta [\tilde{\mathcal{W}} \Lambda^\theta u_j^{\theta\prime}(C_j^\theta) - P_j^{L,\theta}] dC_j^\theta \\ &+ \sum_\theta \sum_j [w_j^{L,\theta} - P_j^{L,\theta} C_j^\theta] dl_j^\theta + \sum_{j,l,k} p_{jl,k} y_{jl,k} \left(\sum_\theta \gamma_{jl,k}^\theta d\ln l_j^\theta + \gamma_{jl,k}^Y d\ln y_{jl,k} \right) \\ =& \tilde{\mathcal{W}} \left[\sum_{i,j,k} p_{ij,k} y_{ij,k} d\ln A_{ij,k} + \sum_\theta \sum_j l_j^\theta [\Lambda^\theta u_j^{\theta\prime}(C_j^\theta) - P_j^\theta] dC_j^\theta \right. \\ &\left. + \sum_\theta \sum_j [w_j^\theta - P_j^\theta C_j^\theta] dl_j^\theta + \sum_{j,l,k} p_{jl,k} y_{jl,k} \left(\sum_\theta \gamma_{jl,k}^\theta d\ln l_j^\theta + \gamma_{jl,k}^Y d\ln y_{jl,k} \right) \right]. \end{aligned} \tag{A.49}$$

Since the term inside the square bracket is identical to what we obtain in equation (A.17), the decomposition remains identical up to the multiplicative constant $\tilde{\mathcal{W}}$.

□

B First-Best Allocation

In this section, we discuss the first-best planning problem, where the Planner can directly specify the allocation for both location choice and consumption. For expositional purposes, we assume that the agglomeration externality depends only on local population size but not on output, i.e. $\gamma_{ij,k}^Y = 0$, as considered in Section 3.3. The problem is given by

$$W = \max_{\{W^\theta, \{C_j^\theta, \mathbf{c}_j^\theta, \mu_j^\theta\}\}, \{\mathbf{x}_{ij,k}, \mathbf{l}_{ij,k}, h_{ij,k}, \mathcal{A}_{ij,k}\}} \mathcal{W}(\{W^\theta\}) \tag{B.1}$$

$$\text{s.t.} \qquad (9), (13)\text{-}(16), \tag{B.2}$$

$$W^\theta = \sum_j \mu_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\mu_j^\theta\}) \tag{B.3}$$

$$\sum_j \mu_j^\theta = 1. \tag{B.4}$$

In the first-best planning problem, the only constraints the Planner faces are resource constraints. Notice that this problem is different from the ones considered in Allen and Arkolakis (2014) and Fajgelbaum and Gaubert (2020), which take the incentive compatibility constraints for location choice as given, which we separately study in Appendix A.4.

The first-order conditions with respect to $c_{ij,k}$, $l^{\theta}_{ij,k}$, $h_{ij,k}$, and $x^{l,m}_{ij,k}$ are

$$P_j^{FB,\theta}\frac{\partial \mathcal{C}_j^{\theta}}{\partial c_{ij,k}}=p^{FB}_{ij,k},\quad p^{FB}_{ij,k}\mathcal{A}_{ij,k}\frac{\partial f_{ij,k}}{\partial l^{\theta}_{ij,k}}=w_i^{FB,\theta},\quad p^{FB}_{ij,k}\mathcal{A}_{ij,k}\frac{\partial f_{ij,k}}{\partial h_{ij,k}}=r_i^{FB},\quad p^{FB}_{ij,k}\mathcal{A}_{ij,k}\frac{\partial f_{ij,k}}{\partial x^{l,m}_{ij,k}}=p^{FB}_{li,m}, \tag{B.5}$$

where $P_j^{FB,\theta}$, $p^{FB}_{ij,k}$, $w_j^{FB,\theta}$, and r_j^{FB} are Lagrange multipliers on constraints (13)-(16). The superscript FB denotes variables from the Planner's first-best allocation. These conditions are identical to the equilibrium conditions (A.12), with the Planner's shadow prices coinciding with equilibrium prices up to a multiplicative constant. Therefore, relative quantities of inputs are not distorted in equilibrium.

The Planner's first-best allocation deviates from the equilibrium when we consider the first-order conditions for C_j^{θ} and μ_j^{θ}. The first-order condition with respect to C_j^{θ} gives us

$$\Lambda^{\theta,FB}u_j^{\theta\prime}(C_j^{FB,\theta})/P_j^{FB,\theta}=1, \tag{B.6}$$

where $\Lambda^{\theta,FB}\equiv\frac{\partial\mathcal{W}}{\partial W^{\theta}}$. That is, the weighted marginal utility of income is equalized across locations for type θ households.

The first-order condition for μ_j^{θ} is

$$\Lambda^{\theta}\left(u_j^{\theta}(C_j^{FB,\theta})-\frac{\partial\psi^{\theta}}{\partial\mu_j^{\theta}}\right)+w_j^{FB,\theta}+\sum_{i,k}p^{FB}_{ji,k}y^{FB}_{ji,k}\frac{\gamma^{\theta}_{ji,k}}{\ell_j^{\theta}}-P_j^{FB,\theta}C_j^{FB,\theta}=\delta^{FB,\theta}/\ell^{\theta}, \tag{B.7}$$

where $\delta^{FB,\theta}$ is the Lagrange multiplier on constraint (B.4).

Now we ask whether the above two conditions can be satisfied in the decentralized equilibrium. Note that $\{\{P_j^{FB,\theta},w_j^{FB,\theta}\},\{p^{FB}_{ij,k}\},r_j^{FB}\}$ coinciding with $\{\{P_j^{\theta},w_j^{\theta}\},\{p_{ij,k}\},r_j\}$ up to a multiplicative constant. Therefore, in order for the decentralized equilibrium to satisfy (B.6), it must be that for some $J^{\theta}>0$,

$$u_j^{\theta\prime}(C_j^{\theta})/P_j^{\theta}=J^{\theta}\quad\text{for all } j,\theta. \tag{B.8}$$

For households' location choice in the decentralized equilibrium (A.5) to satisfy (B.7), agglomeration externalities net of transfers are invariant across locations for a given household type to be equalized:

$$\sum_{i,k}p_{ji,k}y_{ji,k}\frac{\gamma^{\theta}_{ji,k}}{\ell_j^{\theta}}-T_j^{\theta}=H^{\theta}\quad\text{for all } j,\theta \tag{B.9}$$

for some H^{θ}.

We summarize the results as follows.

Proposition B.1. *Consider an economy with $\gamma^{Y}_{ij,k}=0$ for all i,j,k. A decentralized equilibrium solves the problem (B.1) for some welfare weights $\{\Lambda^{\theta}\}$ only if both the marginal utility of income and agglomeration externalities net of transfers are equalized across locations for a given household type (i.e.* (B.8) *and* (B.9) *hold).*

The conditions for optimality are stringent and will not generally be satisfied in a spatial economy. Indeed, there is no reason why the consumption implied by equilibrium prices and transfers set according to (B.9) will equalize marginal utility, since marginal utility never shows up in the equilibrium conditions. For example, consider the case where there are no agglomeration externalities $\gamma^{\theta}_{ij,k} = 0$ and no spatial transfers $T^{\theta}_{j} = 0$. The optimality of the decentralized equilibrium requires $u^{\theta\prime}_{j}((w^{\theta}_{j} + \Pi^{\theta})/P^{\theta}_{j})$ to be equalized across locations. This is generically not satisfied. To see why, suppose that $u^{\theta\prime}_{j}((w^{\theta}_{j} + \Pi^{\theta})/P^{\theta}_{j})$ is equalized across locations in the decentralized equilibrium. Then, we can always perturb the parameters governing marginal utility $u^{\theta\prime}_{j}$ and utility levels u^{θ}_{j} at the same time so that location choices are the same but $u^{\theta\prime}_{j}$ will be different. Such a perturbation leads to dispersion in marginal utility across locations without changing prices or the allocation.

More generally, achieving the first-best requires agents to internalize their technological and fiscal externalities without generating dispersion in marginal utility, but the same argument as before suggests such a condition is knife-edge. To achieve the first-best, the Planner must be able to separately control consumption and the location choice decision, which would require some mechanism to directly influence location choice independent of consumption (i.e. break the incentive compatibility constraint (22)).

C Representative Agent Formulation under GEV Idiosyncratic Location Preferences

In this section, we describe the isomorphic representative agent formulation of location choice under Generalized Extreme Value (GEV) preference shocks.

C.1 Additively Separable Case

Consider an additively separable utility function of the form

$$U^{\theta}_{j}(C^{\theta}_{j}, \epsilon^{\theta}_{j}) = u^{\theta}_{j}(C^{\theta}_{j}) + \epsilon^{\theta}_{j}, \tag{C.1}$$

where ϵ^{θ}_{j} follows a type-I generalized extreme value distribution

$$\mathbb{P}[\epsilon^{\theta}_{1} \leq \bar{\epsilon}_{1}, \ldots, \epsilon^{\theta}_{N} \leq \bar{\epsilon}_{N}] = \exp(-G^{\theta}(\exp(-\nu^{\theta}\bar{\epsilon}_{1}), \ldots, \exp(-\nu^{\theta}\bar{\epsilon}_{N}))), \tag{C.2}$$

where $G^{\theta}(.)$ is a correlation function that is homogeneous of degree one. As is well known since McFadden (1978), this yields the following location choice probability:

$$\mu^{\theta}_{j} = \frac{G^{\theta}_{j}(V^{\theta}_{1}, \ldots, V^{\theta}_{N})V^{\theta}_{j}}{\sum_{i} G^{\theta}_{i}(V^{\theta}_{1}, \ldots, V^{\theta}_{N})V^{\theta}_{i}}, \tag{C.3}$$

where

$$V^{\theta}_{j} \equiv \exp(\nu^{\theta} u^{\theta}_{j}(C^{\theta}_{j})), \quad G^{\theta}_{j} \equiv \frac{\partial G^{\theta}(V^{\theta}_{1}, \ldots, V^{\theta}_{N})}{\partial V^{\theta}_{j}}. \tag{C.4}$$

Now we construct a representative agent formulation that is isomorphic to the above model.

Define a mapping $S_j^\theta(\{\mu_i^\theta\})$ that satisfies the following condition for all j:

$$G_j^\theta(S_1^\theta, \ldots, S_N^\theta)S_j^\theta = \mu_j^\theta. \tag{C.5}$$

The representative agent solves

$$W^\theta = \max_{\{\mu_j^\theta\}_j : \sum_j \mu_j^\theta = 1} \sum_j \mu_j^\theta u_j^\theta(C_j^\theta) - \frac{1}{\nu^\theta} \sum_j \mu_j^\theta \ln S_j^\theta(\{\mu_i^\theta\}). \tag{C.6}$$

The first-order condition with respect to μ_j^θ is given by

$$u_j^\theta(C_j^\theta) - \frac{1}{\nu^\theta} \ln S_j^\theta - \frac{1}{\nu^\theta} \sum_i \mu_i^\theta \frac{\partial \ln S_i^\theta}{\partial \mu_j^\theta} - \delta^\theta = 0, \tag{C.7}$$

where δ^θ is the Lagrange multiplier on the adding up constraint $\sum_j \mu_j^\theta = 1$. Note that

$$\sum_i \mu_i^\theta \frac{\partial \ln S_i^\theta}{\partial \mu_j^\theta} = 1, \tag{C.8}$$

for all j. To see this, we use the fact that $G^\theta(.)$ is homogeneous of degree one and add up (C.5) across j to have $G(S_1^\theta, \ldots, S_N^\theta) = \sum_j \mu_j^\theta$. Taking the derivative with respect to μ_j^θ gives us

$$\sum_i G_i(S_1^\theta, \ldots, S_N^\theta) S_i^\theta \frac{\partial \ln S_i^\theta}{\partial \mu_j^\theta} = 1 \tag{C.9}$$

$$\Leftrightarrow \quad \sum_i \mu_i^\theta \frac{\partial \ln S_i^\theta}{\partial \mu_j^\theta} = 1, \tag{C.10}$$

where we used (C.5) in the second line. Therefore, the first-order condition implies

$$S_j^\theta = \exp(-\nu^\theta \delta^\theta - 1 + \nu^\theta u_j^\theta(C_j^\theta)). \tag{C.11}$$

Thus, $S_j^\theta = \exp(-\nu^\theta \delta^\theta - 1) V_j^\theta$. Combining this with the fact that $G_j^\theta(.)$ is homogeneous of degree zero, we have that (C.5) implies

$$\mu_j^\theta = \exp(-\nu^\theta \delta^\theta - 1) G_j^\theta(V_1^\theta, \ldots, V_N^\theta) V_j^\theta. \tag{C.12}$$

The adding up constraint $\sum_j \mu_j^\theta = 1$ implies that

$$\exp(\nu^\theta \delta^\theta + 1) = \sum_j G_j^\theta(V_1^\theta, \ldots, V_N^\theta) V_j^\theta. \tag{C.13}$$

Therefore we obtain

$$\mu_j = \frac{G_j^\theta(V_1^\theta, \ldots, V_N^\theta) V_j^\theta}{\sum_i G_i^\theta(V_1^\theta, \ldots, V_N^\theta) V_i^\theta}, \tag{C.14}$$

coinciding with the solution to the discrete choice problem (C.20).

Finally, we confirm that indirect utility in the two representations coincides with each other. In the discrete choice problem, indirect utility is given by (see McFadden (1978))

$$W^\theta \equiv \mathbb{E}\left[\max_j \left\{u_j^\theta(C_j^\theta) + \epsilon_j^\theta\right\}\right] \tag{C.15}$$

$$= \frac{1}{\nu^\theta} \ln G^\theta(\exp(\nu^\theta u_1^\theta(C_1^\theta)), \ldots, \exp(\nu^\theta u_N^\theta(C_N^\theta))). \tag{C.16}$$

In the representative agent model, substituting (C.7) and (C.13) into (C.6), we obtain

$$W^\theta = \frac{1}{\nu^\theta} \ln G^\theta(\exp(\nu^\theta u_1^\theta(C_1^\theta)), \ldots, \exp(\nu^\theta u_N^\theta(C_N^\theta))), \tag{C.17}$$

verifying that indirect utility coincides with the original discrete choice formulation.

C.2 Multiplicatively Separable Case

Consider a multiplicatively separable utility function of the form

$$\tilde{u}_j^\theta(C_j^\theta, \epsilon_j^\theta) = \tilde{\epsilon}_j^\theta \tilde{u}_j^\theta(C_j^\theta), \tag{C.18}$$

where $\tilde{\epsilon}_j$ follows a type-II generalized extreme value distribution (multi-variate Fréchet)

$$\mathbb{P}[\tilde{\epsilon}_1^\theta \leq \bar{\epsilon}_1, \ldots, \tilde{\epsilon}_N^\theta \leq \bar{\epsilon}_N] = \exp(-G^\theta((\bar{\epsilon}_1)^{-\nu^\theta}, \ldots, (\bar{\epsilon}_N)^{-\nu^\theta})), \tag{C.19}$$

where $G^\theta(.)$ is a correlation function that is homogeneous of degree one. This yields the following location choice probability:

$$\mu_j^\theta = \frac{G_j^\theta(V_1^\theta, \ldots, V_N^\theta) V_j^\theta}{\sum_i G_i^\theta(V_1^\theta, \ldots, V_N^\theta) V_i^\theta}, \tag{C.20}$$

where

$$V_j^\theta \equiv \tilde{u}_j^\theta(C_j^\theta)^{\nu^\theta}, \quad G_j^\theta(V_1^\theta, \ldots, V_N^\theta) \equiv \frac{\partial G^\theta(V_1^\theta, \ldots, V_N^\theta)}{\partial V_j^\theta}. \tag{C.21}$$

Indirect utility follows

$$\tilde{W}^\theta = G^\theta(V_1^\theta, \ldots, V_N^\theta)^{1/\nu^\theta}. \tag{C.22}$$

Now we construct a representative agent formulation that is isomorphic to the above model. Define the utility function of the representative agent as

$$\mathcal{U}(\{\mu_j^\theta\}) = \sum_j \mu_j^\theta S_j^\theta(\{\mu_i^\theta\})^{-\frac{1}{\nu^\theta}} \tilde{u}_j^\theta(C_j^\theta), \tag{C.23}$$

where $S_j^\theta(\{\mu_i^\theta\})$ is defined, as before, to satisfy the following condition for all j:

$$G_j^\theta(S_1^\theta, \ldots, S_N^\theta) S_j^\theta = \mu_j^\theta. \tag{C.24}$$

The representative agent solves

$$\tilde{W}^\theta = \max_{\{\mu_j^\theta\}_j : \sum_j \mu_j^\theta = 1} \mathcal{U}(\{\mu_j^\theta\}). \tag{C.25}$$

The first-order condition is

$$(S_j^\theta)^{-\frac{1}{\nu^\theta}} \tilde{u}_j^\theta(C_j^\theta) - \frac{1}{\nu^\theta} \sum_i \mu_i^\theta (S_i^\theta)^{-\frac{1}{\nu^\theta}} \tilde{u}_i^\theta(C_i^\theta) \frac{\partial \ln S_i^\theta}{\partial \mu_j^\theta} = \delta^\theta \tag{C.26}$$

where δ^θ is the Lagrange multiplier on the adding up constraint $\sum_j \mu_j^\theta = 1$. Let $x_j^\theta \equiv (S_j^\theta)^{-\frac{1}{\nu^\theta}} \tilde{u}_j^\theta(C_j^\theta)$ and $\boldsymbol{x}^\theta \equiv [x_j^\theta]$. In matrix form, the set of first-order conditions can be expressed as

$$(\mathbf{I} - \boldsymbol{D}^\theta)\boldsymbol{x}^\theta = \delta^\theta \vec{\mathbf{1}}, \tag{C.27}$$

where $\boldsymbol{D}^\theta \equiv [d_{ij}^\theta]$ is an $N \times N$ matrix with $d_{ij}^\theta = \frac{1}{\nu^\theta} \mu_j^\theta \frac{\partial \ln S_j^\theta}{\partial \mu_i^\theta}$, and $\mathbf{I}$ is an $N \times N$ identity matrix. Note that (C.8) implies $\sum_j d_{ij}^\theta = 1/\nu^\theta$. Assuming that $(\mathbf{I} - \boldsymbol{D}^\theta)$ is invertible, the unique solution to (C.27) features

$$x_i^\theta = x^\theta \quad \text{for all } i, \tag{C.28}$$

which in turn implies $(S_j^\theta)^{-\frac{1}{\nu^\theta}} \tilde{u}_j^\theta(C_j^\theta) = K^\theta$ for some constant K^θ. Substituting this expression back into (C.24), we have

$$\mu_j^\theta = (K^\theta)^{-\nu^\theta} G_j^\theta(V_1^\theta, \ldots, V_N^\theta) V_j^\theta. \tag{C.29}$$

Using the adding up constraint $\sum_j \mu_j^\theta = 1$, we can solve for $(K^\theta)^{-\nu^\theta}$ to obtain

$$\mu_j^\theta = \frac{G_j^\theta(V_1^\theta, \ldots, V_N^\theta) V_j^\theta}{\sum_i G_i^\theta(V_1^\theta, \ldots, V_N^\theta) V_i^\theta}, \tag{C.30}$$

as desired. We can plug the above expression into the objective to confirm that the indirect utility also coincides with the original discrete choice formulation:

$$\tilde{W}^\theta = G^\theta(V_1^\theta, \ldots, V_N^\theta)^{1/\nu^\theta}. \tag{C.31}$$

D Alternative Social 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

$$\begin{aligned} dW =& \sum_{i,j,k} p_{ij,k} y_{ij,k} d \ln A_{ij,k} + \sum_{\theta} \sum_{j} \left[\ell^{\theta} \Lambda^{\theta} \frac{\partial \mathcal{U}^{\mathcal{SP},\theta}}{\partial C_j^{\theta}} - l_j^{\theta} P_j^{\theta} \right] dC_j^{\theta} \\ &+ \sum_{\theta} \sum_{j} \ell^{\theta} \Lambda^{\theta} \frac{\partial \mathcal{U}^{\mathcal{SP},\theta}}{\partial \mu_j^{\theta}} d\mu_j^{\theta} + \sum_{\theta} \sum_{j} \left[w_j^{\theta} - P_j^{\theta} C_j^{\theta} \right] dl_j^{\theta} \\ &+ \sum_{\theta} \sum_{j} \sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^{\theta}} \gamma_{jl,k}^{\theta} dl_j^{\theta} + \sum_{j,l,k} p_{jl,k} y_{jl,k} \left(\sum_{\theta} \gamma_{jl,k}^{\theta} d \ln l_j^{\theta} + \gamma_{jl,k}^{Y} d \ln y_{jl,k} \right). \end{aligned} \tag{D.1}$$

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

$$\begin{aligned} & \sum_{\theta} \ell^{\theta} \sum_{j} \left[\Lambda^{\theta} \frac{\partial \mathcal{U}^{\mathcal{SP},\theta}}{\partial C_j^{\theta}} - P_j^{\theta} \mu_j^{\theta} \right] dC_j^{\theta} \\ =& \sum_{\theta} \ell^{\theta} \sum_{j} \mu_j^{\theta} \left[\Lambda^{\theta} \left(\frac{\frac{1}{\mu_j^{\theta}} \frac{\partial \mathcal{U}^{\mathcal{SP},\theta}}{\partial C_j^{\theta}}}{u_j^{\theta\prime}(C_j^{\theta})} - 1 \right) + \Lambda^{\theta} - \frac{P_j^{\theta}}{u_j^{\theta\prime}(C_j^{\theta})} \right] u_j^{\theta\prime}(C_j^{\theta}) dC_j^{\theta} \\ =& \mathbb{E}_{\theta} \left[\mathrm{Cov}_{j|\theta} \left(-\frac{P_j^{\theta}}{u_j^{\theta\prime}(C_j^{\theta})}, u_j^{\theta\prime}(C_j^{\theta}) dC_j^{\theta} \right) \right] + \mathrm{Cov}_{\theta} \left(\Lambda^{\theta} - \mathbb{E}_{j|\theta} \left[\frac{P_j^{\theta}}{u_j^{\theta\prime}(C_j^{\theta})} \right], \mathbb{E}_{j|\theta} \left[u_j^{\theta\prime}(C_j^{\theta}) dC_j^{\theta} \right] \right) \\ &+ \mathbb{E}_{\theta} \left[\Lambda^{\theta} \mathbb{E}_{j|\theta} \left[\left(\frac{1}{\mu_j^{\theta}} \frac{\partial \mathcal{U}^{\mathcal{SP},\theta}}{\partial C_j^{\theta}} - u_j^{\theta\prime}(C_j^{\theta}) \right) dC_j^{\theta} \right] \right]. \end{aligned} \tag{D.2}$$

Consequently, Proposition 1 comes down to

$$dW = \Omega_T + \Omega_{MU} + \Omega_{FE} + \Omega_{TE} + \Omega_R + \Omega_{PM}, \tag{D.3}$$

where

$$\Omega_{PM} = \mathbb{E}_{\theta} \left[\Lambda^{\theta} \mathbb{E}_{j|\theta} \left[\left(\frac{1}{\mu_j^{\theta}} \frac{\partial \mathcal{U}^{\mathcal{SP},\theta}}{\partial C_j^{\theta}} - u_j^{\theta\prime}(C_j^{\theta}) \right) dC_j^{\theta} \right] \right] + \mathbb{E}_{\theta} \left[\Lambda^{\theta} \mathbb{E}_{j|\theta} \left[\frac{\partial \mathcal{U}^{\mathcal{SP},\theta}}{\partial \mu_j^{\theta}} d \ln l_j^{\theta} \right] \right], \tag{D.4}$$

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 also useful in considering welfare criteria and optimal policies that are exclusively based on subcomponents of our decompositions, as in Dávila and Schaab (2025b). Consider the following class of welfare criteria:

$$\mathcal{U}^{SP,\theta}(\{C_j^{\theta}, \mu_j^{\theta}\}) = \sum_{j} (\mu_j^{\theta} + \omega_j^{\theta}) u_j^{\theta}(C_j^{\theta}) - \psi^{\theta}(\{\mu_j^{\theta}\}). \tag{D.5}$$

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

Proposition D.1. *Consider welfare criteria based on* (33) *and* (D.5).

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

$$dW = \Omega_T + \Omega_{MU} + \Omega_{FE} + \Omega_{TE} + \Omega_R + \Omega_{PM}, \tag{D.6}$$

where

$$\Omega_{PM} = \mathbb{E}_\theta\left[\Lambda^\theta \mathbb{E}_{j|\theta}\left[\frac{\omega_j^\theta}{\mu_j^\theta} u_j^{\theta\prime}(C_j^\theta) dC_j^\theta\right]\right] \tag{D.7}$$

with our assumption (D.5). Note that the second term in equation (D.4) 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

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

$$= 0. \tag{D.8}$$

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. Assume that the agglomeration externality depends only on local population size but not on output, i.e. $\gamma_{ij,k}^Y = 0$, as considered in Section 3.3 (or the optimal Pigouvian output subsidy is already implemented, as considered in Appendix A.4.4). In the first case, since Ω_{MU} and Ω_R cancel with Ω_{PM}, optimal transfer policy must satisfy

$$0 = -\sum_i \mu_i^\theta T_i^\theta \frac{\partial \ln \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta} + \sum_i \mu_i^\theta \sum_{l,k} p_{il,k} y_{il,k} \frac{1}{l_i^\theta} \gamma_{il,k}^\theta \frac{\partial \ln \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta}. \tag{D.9}$$

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

$$0 = -\sum_i \frac{\partial \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta}\left[w_i^\theta - P_i^\theta C_i^\theta + \sum_{l,k} p_{il,k} y_{il,k} \frac{1}{l_i^\theta}\gamma_{il,k}^\theta\right]. \tag{D.10}$$

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

$$\mu_j^\theta\left[u_j^{\theta\prime}(C_j^\theta) - P_j^\theta\right] = -\sum_i \frac{\partial \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta}\left[w_i^\theta - P_i^\theta C_i^\theta + \sum_{l,k} p_{il,k} y_{il,k} \frac{1}{l_i^\theta}\gamma_{il,k}^\theta\right], \tag{D.11}$$

which incorporates spatial MU dispersion as well as aggregate efficiency considerations. In the third case, the optimal policy formula is

$$\left(\Lambda^\theta - \mathbb{E}_{j|\theta}\left[\frac{P_j^\theta}{u_j^{\theta\prime}(C_j^\theta)}\right]\right)\mu_j^\theta u^{\theta\prime}(C_j^\theta) = -\sum_i \frac{\partial \hat{\mu}_i^\theta(\boldsymbol{C}^\theta)}{\partial C_j^\theta}\left[w_i^\theta - P_i^\theta C_i^\theta + \sum_{l,k} p_{il,k} y_{il,k} \frac{1}{l_i^\theta}\gamma_{il,k}^\theta\right], \tag{D.12}$$

which incorporates redistribution considerations as well as aggregate efficiency considerations.

E Details on Extensions

E.1 Non-Additive Idiosyncratic Location Preferences

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

We now assume that utility in location i is given by $U_i(C_i^\theta, \epsilon_i^\theta)$. 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

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

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

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

where

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

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 considered here, 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 idiosyncratic preferences.

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

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

where

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

Conditional on the price normalization (28) using this marginal utility, all other terms are unaffected.

E.2 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:

$$U_j^\theta(C_j^\theta, \mathcal{B}_j^\theta, \epsilon_j^\theta) = u_j^\theta(C_j^\theta, \mathcal{B}_j^\theta) + \epsilon_j^\theta, \tag{E.6}$$

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

$$l_i^\theta \mathcal{B}_i^\theta = B_i^\theta g_i^{B,\theta}(\{l_i^{\tilde{\theta}}\}), \qquad \gamma_i^{B,\tilde{\theta}\theta} = \frac{\partial \ln g_i^{B,\theta}}{\partial \ln l_i^{\tilde{\theta}}}, \tag{E.7}$$

where B_i^θ is the fundamental component of the supply 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.[2]

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

[2] Here, we assume that the amenities are rival, i.e. the amenity consumed by each household is the total supply of amenities divided by the population size l_i^θ. For non-rival amenities (e.g. scenery), equation (E.8) involves one more term that is analogous to the amenity externality term.

first-order impact on aggregate welfare can be expressed as

$$
\begin{aligned}
dW = & \underbrace{\sum_{i,j,k} p_{ij,k} y_{ij,k} d\ln A_{ij,k} + \sum_{i,\theta} l_i^\theta \partial_B u_i^\theta \mathcal{B}_i^\theta d\ln B_i^\theta}_{\text{(i) Technology } (\Omega_T)} + \underbrace{\mathbb{E}_\theta \left[\text{Cov}_{j|\theta} \left(-\frac{P_j^\theta}{\partial_C u_j^\theta}, \partial_C u_j^\theta dC_j^\theta \right) \right]}_{\text{(ii) MU Dispersion } (\Omega_{MU})} \\
& + \underbrace{\text{Cov}_{j,\theta}(-T_j^\theta, d\ln l_j^\theta)}_{\text{(iii) Fiscal Externality } (\Omega_{FE})} \\
& + \underbrace{\text{Cov}_{j,\theta} \left(\sum_{l,k} p_{jl,k} y_{jl,k} \frac{1}{l_j^\theta} \gamma_{jl,k}^\theta + \sum_{\tilde{\theta}} \partial_B u_j^{\tilde{\theta}} \mathcal{B}_j^{\tilde{\theta}} \gamma_j^{B,\theta\tilde{\theta}}, d\ln l_j^\theta \right) + \sum_{i,j,k} p_{ij,k} y_{ij,k} \gamma_{ij,k}^Y d\ln y_{ij,k}}_{\text{(iv) Technological Externality } (\Omega_{TE})} \\
& + \underbrace{\text{Cov}_\theta \left(\Lambda^\theta - \mathbb{E}_{j|\theta} \left[\frac{P_j^\theta}{\partial_C u_j^\theta} \right], \mathbb{E}_{j|\theta} \left[\partial_C u_j^\theta dC_j^\theta \right] \right)}_{\text{(v) Redistribution } (\Omega_R)}.
\end{aligned}
\tag{E.8}
$$

where $\partial_B u_j^\theta \equiv \frac{\partial u_j^\theta(C_j^\theta, \mathcal{B}_j^\theta)}{\partial B_j^\theta}$ and $\partial_C u_j^\theta \equiv \frac{\partial u_j^\theta(C_j^\theta, \mathcal{B}_j^\theta)}{\partial C_j^\theta}$. 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 shocks (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 additional 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.

E.3 Isomorphism of Amenity Externality with Idiosyncratic Preferences

In the quantitative spatial equilibrium literature, researchers often argue that amenity congestion externalities are isomorphic to preference shocks and use these specifications interchangeably (e.g. Allen and Arkolakis (2014) for the isomorphism between i.i.d. Fréchet preferences and isoelastic congestion externalities).

At first glance, this isomorphism may appear puzzling in light of our welfare decomposition. The two specifications generate identical location-choice systems and identical marginal-utility profiles. However, as explained in Section 4.2, the amenity-externality specification introduces an additional technological externality term – namely, the covariance between amenity externalities and population changes.

In this section, we discuss how this apparent inconsistency is resolved. Specifically, when amenity externalities take a GEV form – that is, the resulting location choice system exhibits a GEV system – the technological amenity-externality term vanishes.

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

$$U_j(C_j, \mathcal{B}_j) = u_j(C_j) + \mathcal{B}_j, \qquad \mathcal{B}_j = -\frac{1}{\nu} \ln S_j(\{l_i\}), \tag{E.9}$$

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

$$\sum_j l_j \frac{\partial \ln S_j}{\partial l_i} = 1. \tag{E.10}$$

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 $S_j = l_j$, i.e. congestion is iso-elastic to local population size with elasticity ν.

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

$$u_j(C_j) - \frac{1}{\nu} \ln S_j = \bar{u}, \tag{E.11}$$

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,

$$U_j(C_j, \epsilon_j) = u_j(C_j) + \epsilon_j, \tag{E.12}$$

where $\{\epsilon_j\}$ follows specification (38). 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 (E.10). The first-order condition for the representative household problem is

$$u_j(C_j) - \frac{1}{\nu} \ln S_j - \frac{1}{\nu} \underbrace{\sum_i l_i \frac{\partial \ln S_i}{\partial l_j}}_{=1} = \delta, \tag{E.13}$$

which coincides with (E.11) with $\bar{u} = \delta + 1/\nu$. 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 any population reallocation (induced by technological shocks or transfers) cannot alter the aggregate distortion arising from amenity externalities. In particular, the amenity component of the (iv) technological externality term in equation (E.8) comes down to

$$\mathrm{Cov}_j\left(-\frac{1}{\nu} \sum_i l_i \frac{\partial \ln S_i}{\partial \ln l_j}, d \ln l_j\right) = \mathrm{Cov}_j(-\frac{1}{\nu}, d \ln l_j) = 0, \tag{E.14}$$

where we used (E.10) and $\partial_B u_i = 1$. Given that all other terms in equation (E.8) are identical between the two specifications, the aggregate welfare predictions are also isomorphic.

This discussion also clarifies that the isomorphism between models with congestion externality and models with preference shocks following a multivariate Gumbel distribution holds only under the specific functional form of congestion externality given by (E.9) and (E.10). Outside these classes, population reallocation (induced by technological shocks or transfers) alters aggregate distortion arising from amenity externalities, and the isomorphism does not hold in general.[3]

E.4 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 are 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,

$$u_j^\theta(c_j^\theta) = B_j \ln c_j^\theta + D_j, \tag{E.15}$$

where B_j and D_j are slope and intercept parameters specific to the location. 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

$$\max_j u_j^\theta(c_j^\theta) + \epsilon_j^\theta \tag{E.16}$$

$$\text{s.t.} \quad P_j^\theta c_j^\theta = z_j^\theta w_j^\theta (1 + \tau_j^\theta). \tag{E.17}$$

Let

$$C_j^\theta \equiv \frac{w_j^\theta (1 + \tau_j^\theta)}{P_j^\theta} \tag{E.18}$$

denote the consumption of household type θ in location j per efficiency units of labor. With our assumption on the utility function (E.15), we can write the location choice problem as

$$\max_j u_j^\theta(C_j^\theta) + \underbrace{B_j \ln z_j^\theta + \epsilon_j^\theta}_{\equiv \varepsilon_j^\theta}. \tag{E.19}$$

[3] Relatedly, Fajgelbaum and Gaubert (2020) show that, when the amenity externalities multiply the utility function so that $\tilde{U}_j(C_j, \mathcal{B}_j) = \tilde{\mathcal{B}}_j(\{L_i\})\tilde{u}_j(C_j)$, spatial equilibria exhibit misallocation even with iso-elastic amenity externalities, and that spatial transfers can correct this misallocation. As our analyses from Proposition 5 and Appendix C.2 demonstrate, this specification is isomorphic to its log transformation, $U_j(C_j, \mathcal{B}_j) = \mathcal{B}_j(\{L_i\}) + u_j(C_j)$, where $\mathcal{B}_j = \ln \tilde{\mathcal{B}}_j$ and $u_j = \ln \tilde{u}_j$. Consequently, through the lens of our framework, welfare gains in Fajgelbaum and Gaubert (2020) operate by addressing the marginal utility dispersion term (ii). As emphasized above, any population reallocation (induced by technological shocks or transfers) cannot alter the distortion from technological amenity externalities if they take iso-elastic form (i.e. the amenity technology externality term (v) is zero).

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:

$$\max_{\{\mu_j^\theta\}_j:\sum_j \mu_j^\theta=1} \sum_j \mu_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\mu_j^\theta\}). \tag{E.20}$$

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 $\{C_i^\theta\}$:

$$Z_j^\theta(\{C_i^\theta\}) \equiv \mathbb{E}\left[z_j^\theta \middle| j = \arg\max_i u_i^\theta(C_i^\theta) + B_i^\theta \ln z_i^\theta + \epsilon_i^\theta\right]. \tag{E.21}$$

To the extent that the location choice function is invertible,[4] 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:

$$\mathcal{Z}_j^\theta(\{\mu_i^\theta\}) = Z_j^\theta(\{\hat{C}_l^\theta(\{\mu_i^\theta\})\}). \tag{E.22}$$

The goods market clearing conditions are modified as follows

$$\sum_\theta c_{ij,k}^\theta + \sum_{l,m} x_{jl,m}^{i,k} = \mathcal{A}_{ij,k} f_{ij,k}(\mathbf{l}_{ij,k}, \mathbf{x}_{ij,k}) \tag{E.23}$$

$$\mathcal{Z}_j^\theta C_j^\theta \ell^\theta \mu_j^\theta = \mathcal{C}_j^\theta(\mathbf{c}_j^\theta), \tag{E.24}$$

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

$$\sum_{i,k} l_{ji,k}^\theta = \mathcal{Z}_j^\theta \ell^\theta \mu_j^\theta, \tag{E.25}$$

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

$$W = \max_{\{W^\theta,\{C_j^\theta,\mathbf{c}_j^\theta,\mu_j^\theta\}\},\{\mathbf{x}_{ij,k},\mathbf{l}_{ij,k},h_{ij,k},\mathcal{A}_{ij,k}\}} \mathcal{W}(\{W^\theta\}) \tag{E.26}$$

subject to (9), (E.23)-(E.25),

$$W^\theta = \sum_j \mu_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\mu_j^\theta\}) \tag{E.27}$$

$$\{\mu_j^\theta\}_j \in \arg\max_{\{\tilde{\mu}_j\}_j:\sum_j \tilde{\mu}_j=1} \sum_j \tilde{\mu}_j^\theta u_j^\theta(C_j^\theta) - \psi^\theta(\{\tilde{\mu}_j^\theta\}) \tag{E.28}$$

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

$$C_j^\theta = \check{C}_j^\theta \tag{E.29}$$

$$\mathcal{A}_{ij,k} = \check{\mathcal{A}}_{ij,k} \tag{E.30}$$

Applying the envelope theorem, we obtain

$$\begin{aligned} dW &= \Omega_T + \Omega_{MU} + \Omega_{TE} + \Omega_R \\ &+ \sum_\theta \sum_j \left[w_j^\theta - P_j^\theta C_j^\theta \right] \mathcal{Z}_j^\theta \ell^\theta d\mu_j^\theta + \sum_\theta \sum_j \sum_l \left[w_l^\theta - P_l^\theta C_l^\theta \right] \ell^\theta \mu_l^\theta \frac{\partial \mathcal{Z}_l^\theta}{\partial \mu_j^\theta} d\mu_j^\theta \end{aligned}$$

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

$$\begin{aligned} dW &= \Omega_T + \Omega_{MU} + \Omega_{TE} + \Omega_R \\ &+ \underbrace{\mathrm{Cov}_{j,\theta}(-T_j^\theta, d\ln l_j^\theta) + \mathrm{Cov}_{j,\theta}\left(-\sum_l T_l^\theta l_l^\theta \frac{\partial \ln \mathcal{Z}_l^\theta}{\partial \ln \mu_j^\theta}, d\ln l_j^\theta \right)}_{\text{(iii) Fiscal Externality } (\Omega_{FE})}. \end{aligned}$$

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.

E.5 General Spillovers

In our main model, we assumed that agglomeration externalities are purely a function of local population size or producers' output (9). 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)). To capture these general externalities, we extend the spillover function (9) such that

$$\mathcal{A}_{ij,k} = A_{ij,k} g_{ij,k}(\{l_\ell^\theta\}_{\ell,\theta}, \{l_{ij,k}^\theta\}_\theta, y_{ij,k}), \tag{E.31}$$

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

$$\gamma_{ij,k}^{P,\ell\theta} = \frac{\partial \ln g_{ij,k}}{\partial \ln l_\ell^\theta}, \qquad \gamma_{ij,k}^{L,\theta} = \frac{\partial \ln g_{ij,k}}{\partial \ln l_{ij,k}^\theta}, \qquad \gamma_{ij,k}^{Y,\theta} = \frac{\partial \ln g_{ij,k}}{\partial \ln y_{ij,k}}. \tag{E.32}$$

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

$$\Omega_{TE} = \sum_{j,l,k} p_{jl,k} y_{jl,k} \left(\sum_{\ell,\theta} \gamma_{\ell l,k}^{P,\ell\theta} d\ln l_\ell^\theta + \sum_\theta \gamma_{jl,k}^{L,\theta} d\ln l_{jl,,k}^\theta + \gamma_{jl,k}^Y d\ln y_{jl,k} \right). \tag{E.33}$$

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

F Test of Optimality of Spatial Transfers

Our formula in Proposition 2 can also be used to assess whether a given transfer scheme can be rationalized by some SWF. Assuming that the SWF is weakly increasing in the expected utility of each type θ, equilibrium allocations that lead to negative inferred welfare weights cannot be rationalized by any SWF.

Corollary 1. *Consider an economy with $\gamma_{ij,k}^Y = 0$ for all i, j, k. If there exists j and θ such that*

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

then there exists an alternative transfer scheme that improves the welfare relative to the original one for any SWF of the form (17).

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

G Appendix for Application to US Highway Network

G.1 Counterfactual Equilibrium System

To align with the notations used in Allen and Arkolakis (2022), we define the aggregate pre-tax labor income in location j as $Y_j = w_j l_j$ and aggregate post-tax-and-transfer income as $E_j = Y_j + T_j l_j$. Using the results from Online Appendix C.2.1 of Allen and Arkolakis (2022), but leaving the population changes $\hat{l}_j$ in the expression (denoting the proportional change in variable x by $\hat{x} = x'/x$, where x' is the value in the counterfactual equilibrium), we have

$$\hat{A}_i^{-\vartheta}\hat{l}_i^{-\vartheta(\gamma+1)}\hat{Y}_i^{(\vartheta+1)}\hat{P}_i^{-\vartheta} = \left(\frac{E_i}{E_i + \sum_j \Xi_{ij}}\right)\hat{E}_i + \sum_j \left(\frac{\Xi_{ij}}{E_i + \sum_j \Xi_{ij}}\right)\hat{\tilde{t}}_{ij}^{-\frac{\vartheta}{1+\vartheta\lambda}}\hat{P}_i^{-\frac{\vartheta}{1+\vartheta\lambda}}\hat{l}_j^{-\frac{\vartheta(\gamma+1)}{1+\vartheta\lambda}}\hat{Y}_j^{\frac{\vartheta+1}{1+\vartheta\lambda}}\hat{A}_j^{-\frac{\vartheta}{1+\vartheta\lambda}} \tag{G.1}$$

$$\hat{A}_i^{-\vartheta}\hat{l}_i^{-\vartheta(\gamma+1)}\hat{Y}_i^{(\vartheta+1)}\hat{P}_i^{-\vartheta} = \left(\frac{Y_i}{Y_i + \sum_j \Xi_{ji}}\right)\hat{Y}_i + \sum_j \left(\frac{\Xi_{ji}}{Y_i + \sum_j \Xi_{ji}}\right)\hat{\tilde{t}}_{ji}^{-\frac{\vartheta}{1+\vartheta\lambda}}\hat{P}_j^{-\frac{\vartheta}{1+\vartheta\lambda}}\hat{l}_i^{-\frac{\vartheta(\gamma+1)}{1+\vartheta\lambda}}\hat{Y}_i^{\frac{\vartheta+1}{1+\vartheta\lambda}}\hat{A}_i^{-\frac{\vartheta}{1+\vartheta\lambda}} \tag{G.2}$$

In addition, from our location choice system (equation 40),

$$\hat{l}_j = \frac{\exp\left(\nu C_j^{1-\rho}\left(\left(\frac{\hat{E}_j}{\hat{P}_j\hat{l}_j}\right)^{1-\rho} - 1\right)\right)}{\sum_j l_j \exp\left(\nu C_j^{1-\rho}\left(\left(\frac{\hat{E}_j}{\hat{P}_j\hat{l}_j}\right)^{1-\rho} - 1\right)\right)} \tag{G.3}$$

Given baseline values of $\{l_i, E_i, Y_i, C_i, \{\Xi_{ij}\}\}$ and parameters $\{\gamma, \vartheta, \lambda, \rho, \nu, \{\varkappa_j\}\}$, our counterfactuals $\{\hat{l}_i, \hat{P}_i, \hat{Y}_i, \hat{E}_i\}$ are given as the solutions to (G.1), (G.2), (G.3), and the aggregate expenditure and transfers are given by

$$E_j^{'} = Y_j^{'} + T_j^{'}l_j^{'}, \qquad T_j^{'} = \varkappa_j\frac{Y_j^{'}}{L_j^{'}} + T^{*'}, \qquad T^{*'} = -\frac{\sum_j \varkappa_j Y_j^{\prime}}{\sum_j l_j^{\prime}}. \tag{G.4}$$

Given these objects, the changes in leg-specific traffic value are given by

$$\hat{\Xi}_{ij} = \hat{\tilde{t}}_{ij}^{-\frac{\vartheta}{1+\vartheta\lambda}}\hat{P}_i^{-\frac{\vartheta}{1+\vartheta\lambda}}\hat{\Pi}_j^{-\frac{\vartheta}{1+\vartheta\lambda}}, \tag{G.5}$$

where $\hat{\Pi}_j$ is the change in the producer market access, given by

$$\hat{\Pi}_j = \hat{A}_j\hat{l}_j^{\gamma+1}\hat{Y}_j^{-\frac{\vartheta+1}{\vartheta}}. \tag{G.6}$$

G.2 Calibration Details

G.2.1 Estimation of Utility Function Parameters

We will now describe the details of estimating the utility function parameters, using the changes in consumption and population from 1980 to 2000. For estimation purposes, we split our samples into high-skill (4 years of college education or greater) and low-skill (less than 4 years of college education). Denoting t as the year for 1980 and 2000,

$$U_{j,t}(C_{j,t}^{\theta}, \epsilon_{j,t}^{\theta}) = \frac{(C_{j,t}^{\theta})^{1-\rho} - 1}{1-\rho} + B_{j,t}^{\theta} + \epsilon_{j,t}^{\theta}, \tag{G.7}$$

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

$$\mu^{\theta}_{j,t} = \frac{\exp\left(\nu\left[\frac{(C^{\theta}_{j,t})^{1-\rho}-1}{1-\rho} + B^{\theta}_{j,t}\right]\right)}{\sum_i \exp\left(\nu\left[\frac{(C^{\theta}_{i,t})^{1-\rho}-1}{1-\rho} + B^{\theta}_{i,t}\right]\right)}. \tag{G.8}$$

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

$$\Delta\ln(\mu^{\theta}_{j,t}/\mu^{\theta}_{1,t}) = \nu\left[\Delta\frac{(C^{\theta}_{j,t})^{1-\rho}}{1-\rho} - \Delta\frac{(C^{\theta}_{1,t})^{1-\rho}}{1-\rho}\right] + \nu\left[\Delta B^{\theta}_{j,t} - \Delta B^{\theta}_{1,t}\right], \tag{G.9}$$

where $\Delta x_t \equiv x_t - x_{t-1}$ denotes the time-difference for any variable x_t from 1980 to 2000. The identification threat in estimating equation (G.9) is that unobserved location-specific amenity shocks $\Delta B^{\theta}_{j,t}$ 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:

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

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

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

we construct a consistent GMM-estimator of (ρ, ν) that solves

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

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 an additional instrumental variable that interacts $z^{\theta}_{j,t}$ with the baseline consumption: $z^{\theta}_{j,t} \times \ln C^{\theta}_{j,t}$.

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

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

where

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

We run this estimation at the MSA level. We construct real consumption at the MSA level for the years 1980 and 2000, from nominal pre-tax income, Consumer Price Index (CPI), and tax-and-transfer rates. Nominal pre-tax income at the MSA level is derived from the Decennial Censuses for both years. Since the CPI is available at the MSA level after 2008 from Bureau of Economic Analysis (BEA), we back-cast CPI for 1980 and 2000 by applying state-level inflation rates from Hazell et al. (2022) to the 2008 MSA-level CPI, assuming uniform inflation across MSAs within each state. To construct post-tax-and-transfer income, we apply state-level average tax and transfer rates from BEA, again assuming uniformity within each state due to the absence of MSA-level tax payment data before 2008.

G.2.2 Calibration of Transfer Rates

To estimate the rates of spatial transfers $\{\varkappa_j\}$, we use observed pre- and post-tax-and-transfer income. Specifically, we obtain these values at the level of counties in 2012 from Bureau of Economic Analysis (BEA). We then aggregate these values at the level of CBSA. Finally, we compute the ratio between the tax-and-transfer to pre-tax income and denote it by $\varkappa_j^*$. Note that these transfer rates may not necessarily satisfy the government budget constraint $\sum_j T_j l_j = 0$. Therefore, we adjust these rates by a nationwide constant $\overline{\varkappa}$ to exactly satisfy this constraint. Specifically, we set $\varkappa_j = \varkappa_j^* - \overline{\varkappa}$, where $\overline{\varkappa}$ is given by

$$\overline{\varkappa} = \frac{\sum_j \varkappa_j^* w_j l_j}{\sum_j w_j l_j}, \tag{G.16}$$

which satisfies $\sum_j T_j l_j = 0$ with the baseline value of $T^* = 0$ in equation (41).

G.2.3 Calibration of Traffic Values

Denote Ξ_{ij}^* as the average annual daily traffic (AADT) over (i, j) in both directions. We assume that the traffic value is given by

$$\Xi_{ij} = \frac{\gamma_i}{\gamma_i + \gamma_j} \xi \left(\Xi_{ij}^* + \Xi_{ji}^* \right), \tag{G.17}$$

where $\{\xi, \{\gamma_i\}\}$ are parameters, and we normalize $\gamma_1 = 1$. Note that this assumption is consistent with the structural assumption that the link-specific transportation costs are symmetric ($\tilde{t}_{ij} = \tilde{t}_{ji}$ for all i, j).[5] We calibrate $\{\xi, \{\gamma_i\}\}$ by targeting exactly two data moments. First, transfers (trade

[5]To see this, start with an assumption that $\Xi_{ij} = \vartheta_{ij} \xi \left(\Xi_{ij}^* + \Xi_{ji}^* \right)$. By assuming that the value of traffic in both ways is proportional to AADT, we have $\Xi_{ji} = (1 - \vartheta_{ij}) \xi \left(\Xi_{ij}^* + \Xi_{ji}^* \right)$. Furthermore, from the gravity equation of traffic (equation 27) of Allen and Arkolakis (2022) together with the symmetric transportation costs assumption, Ξ_{ij}/Ξ_{ji} can be written as γ_i/γ_j. Combining these two expressions, we obtain the expression (G.17).

surplus) are consistent with the net surplus in traffic inflows to outflows, such that

$$\sum_{i'} \Xi_{i'j} - \sum_{k} \Xi_{jk} = E_j - Y_j, \tag{G.18}$$

where Y_j is aggregate pre-tax labor income in location j, and E_j is aggregate post-tax-and-transfer income in location j. Second, we set the average model-implied share of consumption expenditure over goods produced in other locations to coincide with the share of tradables in the consumption basket of the US (27.6 percent; Johnson 2017). That is,

$$\frac{\sum_{i \neq j} X_{ij}}{\sum_i Y_i} = 0.276,$$

where X_{ij} are values of goods that are produced in i and consumed in j, whose expression is given by equation (34) of Allen and Arkolakis (2022). Given that we have N parameters with N equations (note that one equation of (G.18) is redundant), we can exactly match these two sets of moments.

Through this procedure, we find that $\sum_{i,j:i\neq j} \Xi_{ij}/(\sum_i Y_i) = 0.726$. This value is somewhat smaller than the value one, assumed by Allen and Arkolakis (2022).

G.3 Additional Tables

Table G.3.1: Unpacking Deviations from the Δ Technology Term: Mean and Standard Deviation

Variable	Mean	Std. Dev.
$(\Delta W - \Omega_T)/\Omega_T$	-0.54	0.24
Ω_{MU}/Ω_T	0.17	0.18
Ω_{FE}/Ω_T	-0.02	0.04
Ω_{TE}/Ω_T	-0.70	0.19
$\Omega_{TE,S}/\Omega_T$	-0.68	0.19
$\Omega_{TE,A}/\Omega_T$	-0.02	0.02
Residual	0.01	0.00

Note: This table presents the mean and standard deviation of each component of welfare gains in Proposition 1 as a ratio of the Δ technology term (Ω_T) across 704 link improvement simulations. We further decompose Ω_{TE} into the shipment congestion externality $\Omega_{TE,S}$ and agglomeration externality $\Omega_{TE,A}$ using the formula (49). Residual is defined as $\Delta W - (\Omega_T + \Omega_{MU} + \Omega_{FE} + \Omega_{TE})$ divided by Ω_T.

G.4 Sensitivity Analysis to Alternative Specifications

In this section, we assess the sensitivity of our results from Section 5.3. In that section, we assess the heterogeneous welfare gains from transportation infrastructure improvement by reducing the exogenous component of shipment costs $\tilde{t}_{ij} = \tilde{t}_{ji}$ by 1 percent for each of the 704 links. We decompose the resulting welfare effects using Proposition 1 under alternative model specifications. For each specification, we present a version of Figure 2, where the top panel of each figure shows the distribution of deviations in welfare gains from the technology term, measured by $(\Delta W - \Omega_T)/\Omega_T$, while the bottom panel displays the contributions of the three key components: changes in MU dispersion (Ω_{MU}/Ω_T), fiscal externality (Ω_{FE}/Ω_T), and technological externality (Ω_{TE}/Ω_T).

Figure G.4.1 replaces our baseline CRRA utility (with $\rho = 1.90$) with log utility ($\rho \to 1$), consistent with AA's specification. We also adopt AA's baseline calibration of $\nu = 3.3$.[6] Under this specification, the contributions of the three components remain qualitatively similar to the baseline, with a somewhat smaller role for the MU dispersion term (mean of 0.08 and standard deviation of 0.10 log points, compared to 0.17 and 0.18 in the baseline). This is consistent with the reduced dispersion in marginal utility under log utility. The technological and fiscal externality terms are virtually unchanged from the baseline.

Figure G.4.2 further removes transfers, yielding a specification identical to AA's original model. As expected, the fiscal externality term (Ω_{FE}) is mechanically zero and does not contribute to deviations from the technology term. However, the MU dispersion and technological externality terms remain similar to those in Figure G.4.1. Notably, MU dispersion continues to play a significant role in welfare deviations, in addition to technological externalities, which is emphasized by AA as the primary source of heterogeneity in welfare gains.

Figure G.4.3 reverts to our baseline calibration of ρ and ν, but instead considers an alternative transfer rule. Rather than setting transfers based on location-specific tax rates and lump-sum components (as in equation 41), we consider a rule where each location receives a lump-sum transfer (per capita):

$$T_j = \varkappa_j + T^*, \tag{G.19}$$

where $\varkappa_j$ denotes a fixed per capita transfer, and $T^* = -\sum_j \varkappa_j l_j$ ensures the government budget constraint is satisfied, i.e. $\sum_j T_j l_j = 0$. We calibrate $\varkappa_j$ to match the observed distribution of post-tax-and-transfer income, as is similar to our baseline calibration (see Appendix G.2.2). Under this alternative transfer rule, we again observe broadly similar patterns.

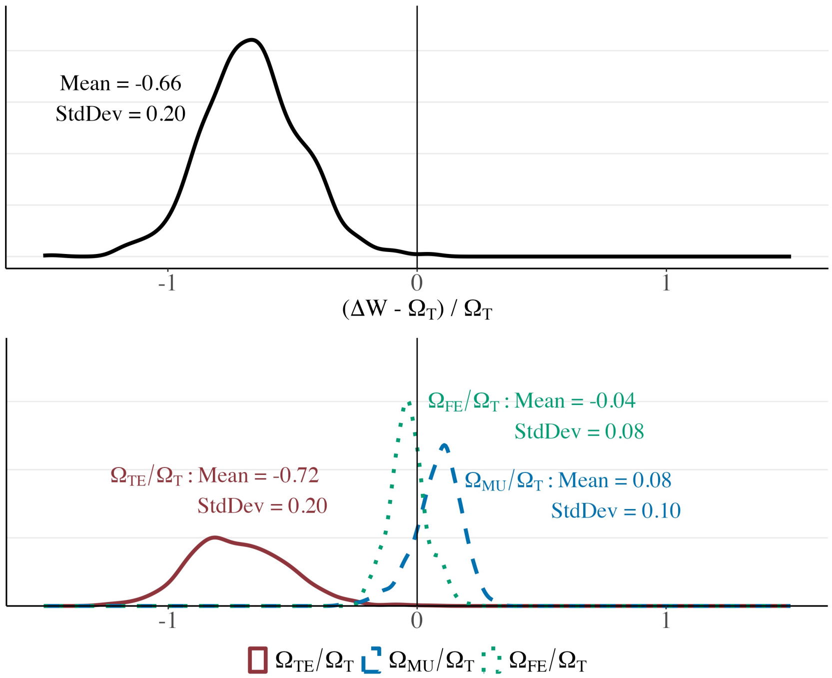
Figure G.4.4 returns to our baseline calibration of ρ and ν, including transfers, but increases the trade elasticity from $\vartheta = 8$ to $\vartheta = 16$. The results are qualitatively unchanged, with two slight changes. First, the mean of the technological externality term increases modestly to −0.81 log points, compared to −0.70 in the baseline. Second, both the mean and dispersion of the marginal utility dispersion and fiscal externality terms slightly decline. The increase in the technological externality is consistent with the interpretation that a higher ϑ amplifies the reallocation of trade flows, thereby intensifying the changes in shipment congestion externalities. The decline in the other terms suggests that a higher ϑ allows trade flows to absorb shocks more flexibly, without inducing large population reallocations to restore equilibrium.

Figure G.4.5 presents results under a lower congestion externality elasticity, where λ is reduced by half, from 0.092 to 0.046. As expected, the contribution of the technological externality term declines. In contrast, the MU dispersion and fiscal externality terms increase. This shift reflects the fact that weaker congestion externalities increase the price effects of technological shocks, leading to greater population reallocation and, consequently, stronger changes in MU

[6] As noted in the paper, with $\rho \to 1$, our utility specification in equation (40) is positively and normatively isomorphic to AA's original specification, which takes a multiplicative function of C_j and an isoelastic negative amenity externality with elasticity $-1/\nu$, with degenerate idiosyncratic preferences (Sections 4.1 and 4.3). AA's calibration of $1/\nu = 0.3$ implies $\nu \approx 3.3$.

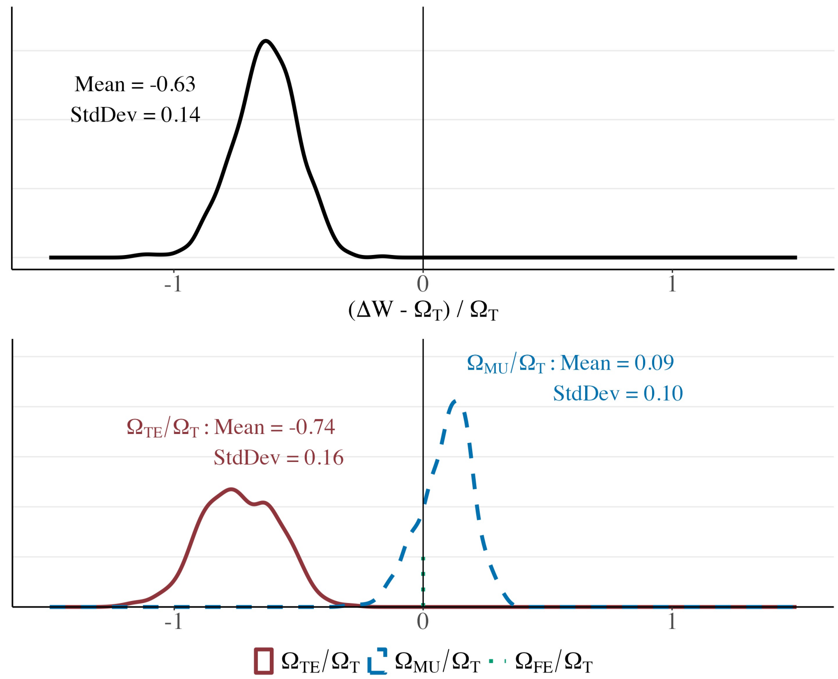
dispersion and fiscal externalities.

Figure G.4.1: Unpacking Deviations from the Δ Technology Term: log Utility



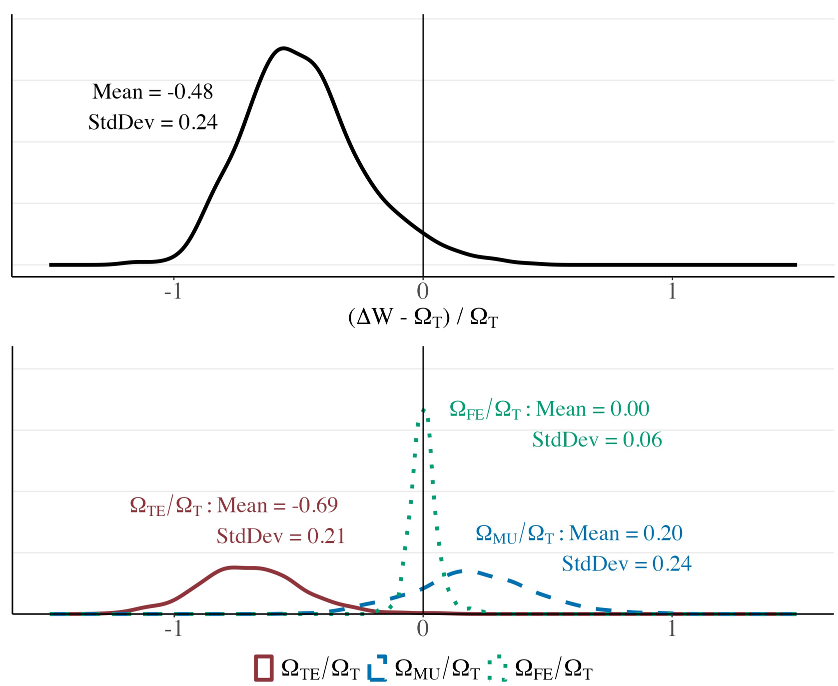
Note: The figure is the counterpart of Figure 2, where we use the same utility parametrization of Allen and Arkolakis (2022), i.e. $\rho \to 1$ (log utility) and $\nu = 3.3$.

Figure G.4.2: Unpacking Deviations from the Δ Technology Term: log Utility & No Transfer



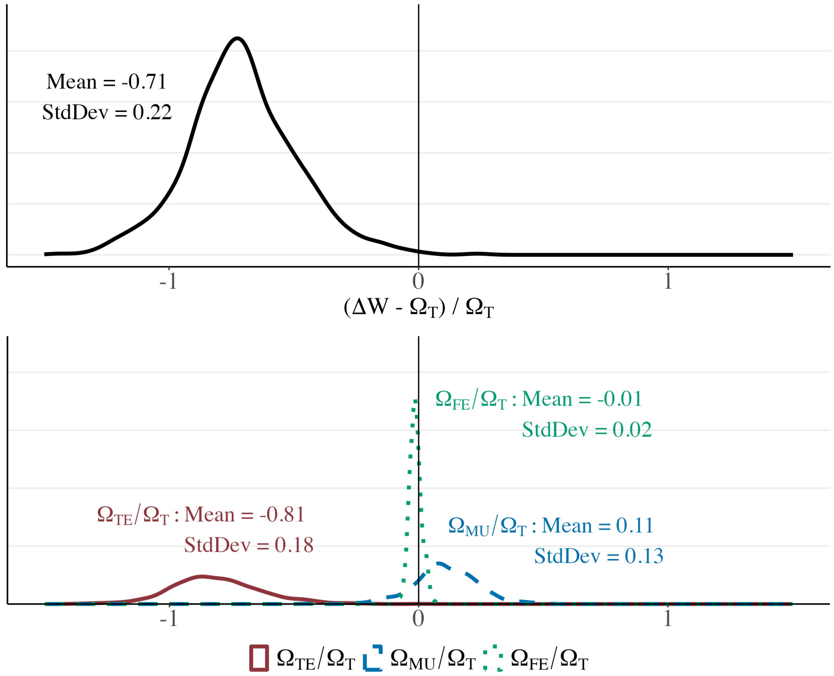
Note: The figure is the counterpart of Figure 2, where we use the same utility specification and parametrization of Allen and Arkolakis (2022), i.e. $\rho \to 1$ (log utility) and $\nu = 3.3$, and assuming no transfers.

Figure G.4.3: Unpacking Deviations from the Δ Technology Term: Alternative Transfer Rule (Lump-Sum)



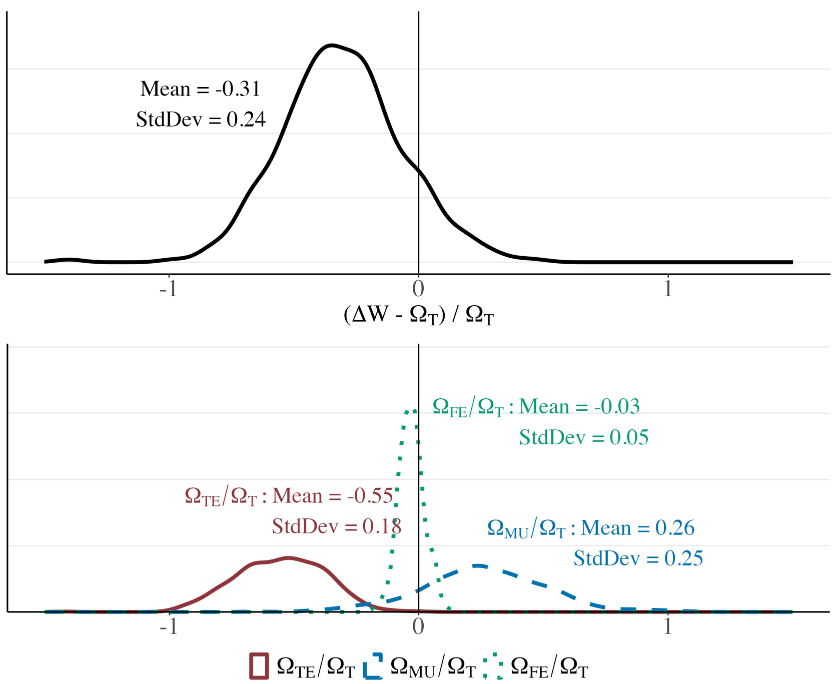
Note: The figure is the counterpart of Figure 2, where we use an alternative rule of spatial transfers based on lump-sum transfers, instead of proportional income tax.

Figure G.4.4: Unpacking Deviations from the Δ Technology Term: Higher ϑ



Note: The figure is the counterpart of Figure 2, where we use our baseline model and set a higher trade elasticity $\vartheta = 16$, instead of $\vartheta = 8$ in our baseline specification.

Figure G.4.5: Unpacking Deviations from the Δ Technology Term: Lower λ



Note: The figure is the counterpart of Figure 2, where we use our baseline model and set a lower congestion elasticity $\lambda = 0.046$, instead of $\lambda = 0.092$ in our baseline specification.

H Application: Local Productivity Shocks

In this Appendix, we present an alternative application of the model introduced in Section 5, where we consider location-specific productivity shocks instead of changes in shipment costs. Specifically, we simulate a 1% increase in the exogenous productivity component A_i for each of the 228 nodes (CBSAs) in the highway network.

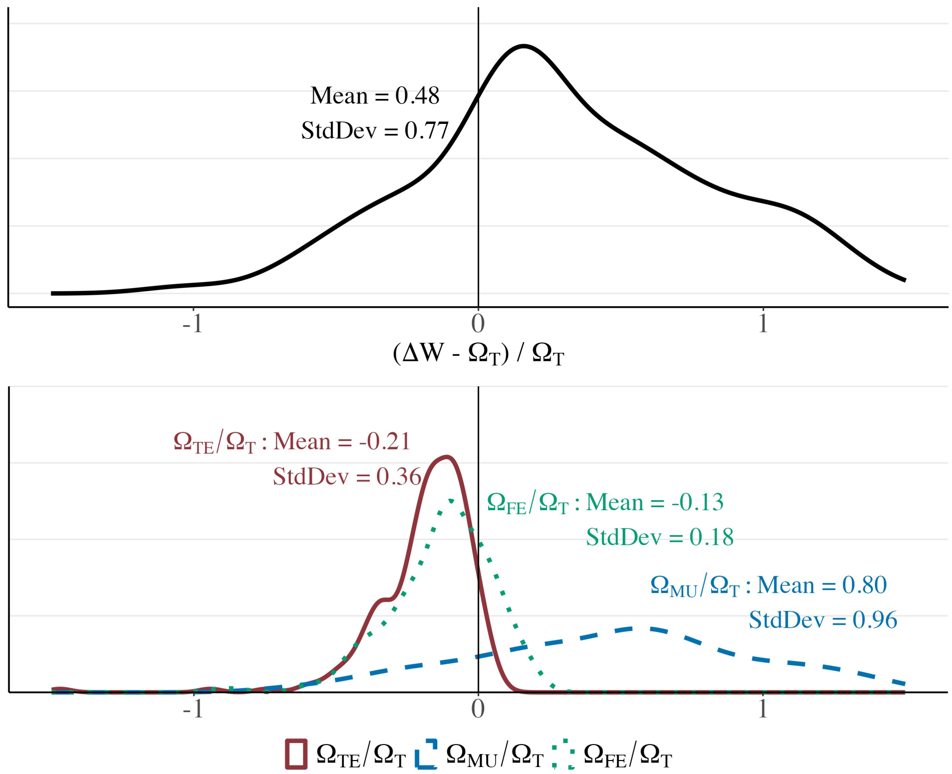
The top panel of Figure H.0.1 displays the distribution of welfare deviations from the technology term, measured by $(\Delta W - \Omega_T)/\Omega_T$, analogous to Figure 2 in the main text. We find a substantial deviation, with a mean of 0.48 log points and a standard deviation of 0.77.

The bottom panel decomposes the welfare deviations into three key components: changes in MU dispersion (Ω_{MU}/Ω_T), fiscal externality (Ω_{FE}/Ω_T), and technological externality (Ω_{TE}/Ω_T). Compared to the transportation counterfactuals, the relative contributions differ. The technological externality is on average negative (mean −0.21 log points; standard deviation 0.36). Table H.0.1 shows that both shipment congestion and location-specific agglomeration externalities contribute roughly equally to the mean, with greater dispersion attributed to the congestion externality.

The MU dispersion term is on average positive and sizable (mean 0.80 log points; standard deviation 0.96). The fiscal externality also plays a significant role, with a mean of −0.13 and a standard deviation of 0.18 log points.

Overall, relative to the transportation improvement counterfactual, deviations from the technology term are more strongly driven by changes in MU dispersion and fiscal externalities in this exercise.

Figure H.0.1: Unpacking Deviations from the Δ Technology Term: Local Productivity Shocks



Note: The figure is the counterpart of Figure 2, where we consider local productivity improvement to each location, instead of infrastructure improvement.

Table H.0.1: Unpacking Deviations from the Δ Technology Term: Local Productivity Shocks

Variable	Mean	Std. Dev.
$(\Delta W - \Omega_T)/\Omega_T$	0.48	0.77
Ω_{MU}/Ω_T	0.80	0.96
Ω_{FE}/Ω_T	-0.13	0.18
Ω_{TE}/Ω_T	-0.21	0.36
$\Omega_{TE,S}/\Omega_T$	-0.11	0.31
$\Omega_{TE,A}/\Omega_T$	-0.10	0.11
Residual	0.01	0.08

Note: This table presents the mean and variance of each component of welfare gains in Proposition 1 as a ratio of the Δ technology term (Ω_T) across 228 simulations, where each simulation increases the productivity of each of 228 nodes (CBSAs) by one percent. We further decompose Ω_{TE} into the shipment congestion externality $\Omega_{TE,S}$ and agglomeration externality $\Omega_{TE,A}$ using the formula (49). Residual is defined as $\Delta W - (\Omega_T + \Omega_{MU} + \Omega_{FE} + \Omega_{TE})$ divided by Ω_T.

I Application: Ex-post Welfare Changes from Regional Growth in the US

We further demonstrate the usefulness of our formula through another application with an "ex-post" approach. We study the welfare changes in the US economy implied by the observed changes in spatial allocations during 2010-2019.

I.1 Bringing the Formula to Data

We assume that preferences satisfy $u_{j,t}^{\theta}(C_{j,t}^{\theta}) + \epsilon_{j,t}^{\theta} = \frac{(C_{j,t}^{\theta})^{1-\rho}}{1-\rho} + B_{j,t}^{\theta} + \epsilon_{j,t}^{\theta}$ where $\epsilon_{j,t}^{\theta}$ follows a type-I extreme value distribution with scale parameter ν, and t indexes time.

For any type θ at any location i at time t, suppose we observe population $l_{j,t}^{\theta}$, pre-tax income $w_{j,t}^{\theta} + \Pi_{t}^{\theta}$, net transfers T_{j}^{θ}, and price indices $P_{j,t}^{\theta}$ as well as total sales in each location $\sum_{l,k} p_{jl,k} y_{jl,k}$. Furthermore, suppose we know the utility function parameters $\{\rho, \nu\}$ and agglomeration functions $\{g_{ij,k}(\cdot)\}$. Then, we can use Proposition 1 to back out the first-order welfare changes between any dates that are attributable to the (ii) Δ MU dispersion (Ω_{MU}), (iii) Δ fiscal externality (Ω_{FE}), and (iv) Δ technological externality (Ω_{TE}) terms.

Now we show that the (i) Δ technology term (Ω_{T}) can be read directly from the data. Let GDP_i be the nominal GDP of MSA i and Y_i be the 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:

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

Furthermore, assuming that agglomeration externality depends only on the local population size (but not output) for this application for simplicity, we have

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

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

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

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

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

By the definition of real GDP,

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

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

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

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.

Finally, if one is willing to take a stance on welfare weights across different households $\{\Lambda^\theta\}$, then the first-order welfare changes that are attributable to the (v) redistribution term (Ω_R) are recovered as well.

Data. We implement the above approach in the context of Metropolitan Statistical Areas (MSAs) in the United States during the period 2010-2019. Our sample consists of 214 MSAs. Following Diamond (2016) and Fajgelbaum and Gaubert (2020), we consider two ex-ante types based on educational attainments: high-skill (4 years of college education or greater) and low-skill (less than 4 years of college education). We also restrict our analysis to the working-age (age 18-64) population. We construct our dataset using a combination of the BEA regional economic accounts, American Community Survey (ACS) (IPUMS-USA, Ruggles, Flood, Goeken, Grover, Meyer, Pacas, and Sobek 2023), March supplement of the Current Population Survey (March CPS) (IPUMS-CPS, Flood, King, Rodgers, Ruggles, Warren, and Westberry 2023), Consumer Expenditure Survey (CEX), and IRS (Internal Revenue Office) Statistics of Income (SOI) data.

We first obtain population, pre-tax income, and transfer receipts by MSAs from the BEA. We allocate them to each skill group based on their shares in each MSA using 5-year samples of the ACS. We obtain tax payments by county from the IRS SOI and then aggregate them to the MSA level using the crosswalk provided by the NBER. We further allocate them to each skill group based on the aggregate shares in tax payments by each skill group using the March CPS. Net transfers $T_{j,t}^\theta$ are constructed as the difference between transfer receipts and tax payments, but we adjust them by adding a common constant so as to ensure government budget balance. Finally, we construct price indexes for each MSA and skill group as follows. The BEA provides price indexes for four broad categories at the MSA level: goods, housing, utilities, and other services. We compute the expenditure share on these four categories for each skill group using the CEX. We then construct price indexes for each MSA and skill group by weighting the price level of the four categories using the expenditure weight for each skill group. With pre-tax income, $w_{j,t}^\theta + \Pi_t^\theta$, net transfers $T_{j,t}^\theta$, and price indexes $P_{j,t}^\theta$, we compute consumption for each location-type as $C_{j,t}^\theta = \frac{w_{j,t}^\theta + \Pi_t^\theta + T_{j,t}^\theta}{P_{j,t}^\theta}$.

Table I.2.1: Welfare Changes in the US 2010-2019

	dW	Ω_T	Ω_{MU}	Ω_{FE}	Ω_{TE}	Ω_R
2010-2015	2.71%	2.42%	0.12%	0.02%	0.01%	0.15%
2015-2019	2.17%	1.52%	0.03%	0.01%	-0.07%	0.69%

Note: This table reports welfare decompositions based on Proposition 1 for observed changes in spatial allocations across US MSAs during the periods of each row. Welfare is expressed in units of efficiency labor equivalent in the initial period. All reported numbers are annualized changes.

Parameterization. We use the same estimates of (ρ, ν) as in Table 2. We assume away input-output linkages and set the total sales of each skill type in each location as their pre-tax income. We abstract from output-based externalities ($\gamma^Y_{ij,k} = 0$), but instead introduce population-based productivity externalities among and between high and low skills. Specifically, we follow Fajgelbaum and Gaubert (2020) by setting them to $(\gamma^l_{ij,l}, \gamma^h_{ij,l}, \gamma^l_{ij,h}, \gamma^h_{ij,h}) = (0.003, 0.044, 0.02, 0.053)$, where $\gamma^\theta_{ij,\theta'}$ corresponds to the productivity spillover from type θ to θ' for the goods shipped from i to j. The skill type h denotes high-skill and l denotes low-skill. For welfare weights, we assume utilitarian welfare: $\Lambda^\theta = 1$ for all θ.

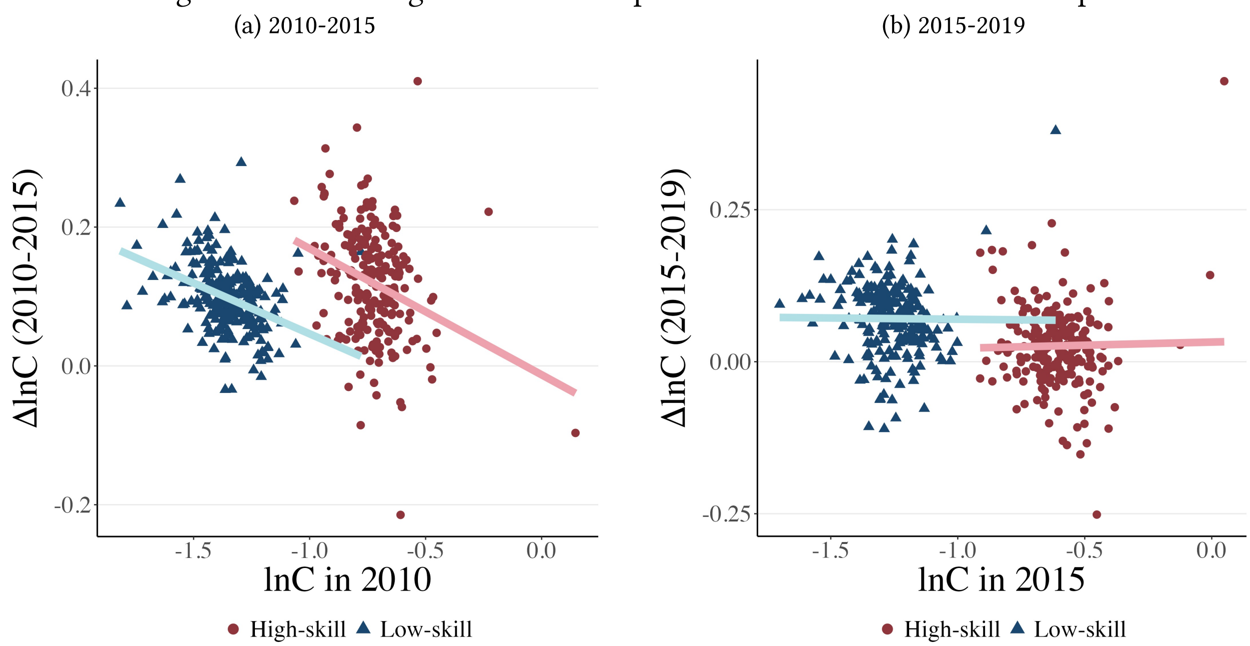
I.2 Results

Table I.2.1 shows welfare decompositions based on observed changes in spatial allocations during the periods 2010-2015 and 2015-2019. All of these numbers are annualized changes and expressed in units of efficiency labor equivalent in the initial period. Combined with our choice of numeraire (28), these numbers answer the following question: "If we were to achieve the same welfare change by uniformly increasing utility in all locations for all household types, what percent increase in uniform labor productivity would we need?"

Between 2010 and 2015, aggregate welfare increased by an efficiency labor equivalent of 2.71%. The largest contributor is the (i) Δ technology term, but Hulten's characterization understates the welfare gain of this period. The (ii) Δ MU Dispersion term plays a non-trivial secondary role in raising welfare. This is driven by the reduction in spatial consumption inequality within skill groups, as shown by the left panel of Figure I.2.1. The (v) Δ redistribution terms also contribute to some extent, reflecting the catch-up of the low-skill group to the high-skill group. The (iii) Δ fiscal externality and (iv) Δ technological externality terms also contribute to the welfare gain, but their magnitudes are small.

In 2015-2019, aggregate welfare increased by an efficiency labor equivalent of approximately 2.17%. The (i) Δ technology term again understates the welfare gain, but now, the (ii) Δ MU Dispersion term is close to zero. Instead, the (v) Δ redistribution term plays a substantial role in the welfare gain. The right panel of Figure I.2.1 explains why. During this period, the *within* group spatial distribution of consumption stopped converging, but there was convergence of consumption *across* skill groups.

Figure I.2.1: Convergence in Consumption Across MSAs and Skill Groups



Note: The figure plots consumption changes for each MSA and skill group against their initial consumption level. The solid line is the best linear fit within each skill group.