Welfare Analysis with Discrete Choice in General Equilibrium

David R. Baqaee
UCLA

Ariel Burstein
UCLA*

October 19, 2026

Abstract

Discrete-choice general-equilibrium models capture rich heterogeneity in where people live and work, but this heterogeneity complicates the aggregation of welfare gains and losses. We measure aggregate economic surplus by the fraction of endowments that can be saved after a change while keeping every household as well off as before. This is a general-equilibrium counterpart to cost-benefit analysis that allows prices, wages, and choices to adjust as winners compensate losers. It requires neither utility cardinalizations nor Pareto weights. Under perfect competition, the first-order welfare elasticity with respect to a producer's productivity equals its sales relative to GDP, regardless of the distribution of preferences and technologies. Beyond first order, we provide exact characterizations using supply and demand functions and second-order approximations using observable income and expenditure shares and supply and demand elasticities. We compare this measure with commonly-used alternative measures of aggregate welfare like real GDP, the sum of compensating variations, and average utility. Using a version of the Allen and Arkolakis (2014) spatial model, we show that welfare estimates and policy prescriptions based on average-utility vary erratically across observationally equivalent cardinal representations of the same ordinal preferences. Our welfare measure is invariant to these arbitrary assumptions.

*We thank Rodrigo Adao, Fernando Alvarez, John Asker, Lorenzo Caliendo, Pablo Fajgelbaum, Hannes Malmberg, Yuhei Miyauchi, Will Rafey, Stephen Redding, Jon Vogel, Pierre-Olivier Weill for their comments. We are especially grateful to Natalie Bau for her insights.

1 Introduction

General-equilibrium models in which agents make discrete decisions about where to live and work are now standard in trade, spatial, labor, and macroeconomics. These models are popular because they capture rich heterogeneity in preferences, skills, and individual outcomes. Yet they make it surprisingly hard to answer a basic question: what is the aggregate welfare value of a change in the economy when agents are affected differently by that change? A meaningful answer is important for summarizing the overall welfare consequences of economic change and for guiding policy.

The most commonly-used measure in the spatial literature is the cross-sectional average of individual utilities.[1] A fundamental issue with this measure is that it is not invariant to monotone transformations of individual utility functions. Replacing a utility function with a monotone transformation that generates exactly the same choice behavior can nonetheless alter the social welfare ranking of allocations. For instance, in the spatial literature, a standard utility specification is

$$u_h(c, r) = c\varepsilon_{hr},$$

where c is consumption and ε_{hr} is h's idiosyncratic taste for choosing location r. This specification is observationally equivalent to

$$\tilde{u}_h(c, r) = c\varepsilon_{hr} / \max_{r'}\{\varepsilon_{hr'}\},$$

which normalizes each agent's highest location-specific taste to one. Although both specifications imply the same individual rankings, and hence choices, they imply different rankings of social allocations and different optimal policies. The issue is that there is no canonical basis for choosing between these normalizations. Normalizing tastes so that every agent assigns a taste of one to their favorite location is just as plausible as allowing some agents to have higher tastes in their favorite location than others. The data do not distinguish between these cardinalizations, and the resulting aggregate welfare measure hence depends on untestable and nonintuitive assumptions about the interpersonal scale of utility.[2]

[1] Examples include, among many others, Redding (2016), Redding and Rossi-Hansberg (2017), Fajgelbaum et al. (2019), Dingel and Tintelnot (2020), Allen and Arkolakis (2022), Monte et al. (2023), and Tsivanidis (2026). Average utility is also widely used outside spatial models; see, for example, Kleiner and Soltas (2023) in a model of occupational choice.

[2] See Section 6.3 for a discussion and worked-out example. Although this measure is sometimes called "expected utility," average utility is not a von Neumann–Morgenstern expected utility function. Expected utility functions represent ordinal preference relations over lotteries; they are not expectations of utility

In this paper, we develop a different approach for measuring aggregate welfare changes in discrete choice economies. Following Baqaee and Burstein (2025a), and inspired by Debreu (1951), we measure the aggregate surplus created by a change in primitives by the fraction of factor endowments left over if winners from that change were to compensate the losers. This measure has the following attractive properties. First, because compensation is defined by households' indifference conditions, the measure depends only on ordinal preferences and is invariant to arbitrary cardinalizations of utility. Second, it respects Pareto improvements: when every household is better off, the measure rises. Third, it is distribution-neutral: it quantifies aggregate surplus without specifying Pareto weights or deciding which households ought to receive the surplus. Our contribution is to introduce this notion of aggregate economic surplus into general equilibrium economies with discrete choice, characterize it in terms of observable supply and demand curves, and show how its implications differ from those of standard aggregate welfare measures.

Our approach can be viewed as an extension of traditional cost-benefit analysis to general equilibrium settings. Traditionally, cost-benefit analysis measures the surplus from a change via the sum of compensating variations (see Harberger, 1971).[3,4] The sum of compensating variations measures the "money" left over after winners compensate losers, holding relative prices constant. Small and Rosen (1981) characterize this welfare measure in partial equilibrium discrete-choice models and link it to compensated choice probabilities. The sum of compensating variations is unsuitable for general equilibrium analyses because relative prices are endogenous: a change that looks like a potential Pareto-improvement holding prices constant may fail to be one once we account for the fact that prices would adjust in the event of an attempted compensation. This issue becomes particularly severe for optimal policy problems where policies that maximize the sum of compensating variations will often have bang-bang, and frequently inefficient, corner solutions.[5]

We resolve this problem by determining compensating transfers, prices, wages, and

values across households with different preferences.

[3]The pioneers in the discrete choice literature, like McFadden (1981) and Anderson et al. (1992), emphasize that consumer welfare in discrete choice models should be measured using compensating or equivalent variations of the individual agents, as in standard consumer theory. Dagsvik and Karlström (2005) show how to calculate the distribution of compensating variation across agents using compensated choice probabilities.

[4]While this approach is common in the industrial organization literature, it is less common in general equilibrium spatial and occupational choice models. One exception is Kim and Vogel (2020), who study a model where workers choose among a discrete number of sectors including non-employment. They show that the elasticity of the sum of compensating variations to wages is the income share of each sector.

[5]Intuitively, if the policymaker takes all resources away from some households, the relative price of what those households like to consume will fall dramatically, which makes it possible to compensate them cheaply at the new prices with money left over. We provide explicit examples in Section 6.

choices jointly in general equilibrium. We then measure aggregate surplus by the fraction of factor endowments, rather than money, that can be saved while keeping every household at its status-quo welfare. Although our definition follows Baqaee and Burstein (2025a), their characterization does not carry over to our discrete-choice setting. We therefore develop a different characterization of aggregate welfare that does not rely on their fictional representative-agent construction.

The structure of the paper is as follows. In Section 2, we set up preferences, technologies, and define the perfectly competitive equilibrium. Although the equilibrium is perfectly competitive, the model is otherwise fairly flexible: agents can differ flexibly in their skills and preferences over locations and occupations. Production uses different types of primary factors and intermediate inputs, and accommodates input-output networks and costly trade. Market imperfections and externalities are dealt with later in the paper.

In Section 3, we define our aggregate welfare statistic and show that it can be characterized using compensated supply and demand functions. Compensated supply and demand functions map relative wages and prices to labor supply and spending choices across locations, holding each agent on a fixed indifference curve. The change in aggregate welfare from a shock is the adjustment in total factor productivity that makes the net change in the sum of consumer and producer surplus — the areas under compensated demand and above compensated supply curves — zero. We show how to obtain these compensated supply and demand functions using either household-level simulations or, building on an insight from Bhattacharya (2015, 2018), using market-level uncompensated supply and demand.

In Section 4,we generalize Hulten's theorem to perfectly competitive economies with discrete choice: observed sales shares are sufficient statistics for the first-order welfare effects of productivity shocks. Specifically, the welfare elasticity with respect to a producer's productivity equals its initial sales relative to GDP, regardless of preferences and technologies. Thus, with our definition, aggregate welfare obeys Hulten's theorem just as in perfectly competitive representative-agent spatial economies (Donaldson, 2025). Intuitively, to a first order, one does not need to solve for how households switch locations or occupations in response to a shock because when a household switches choices, the pecuniary gain from switching is exactly offset at the margin by the amenity loss required to make the household willing to switch. Beyond first order, we derive an exact formula that accumulates productivity changes weighted by sales relative to GDP along the compensated equilibrium, and second-order approximations using initial expenditure shares and supply and demand elasticities.

In Section 5, we discuss extensions to our basic framework. We show how to account

for market imperfections and externalities. We also discuss how to relax the assumption that compensations can be implemented through individual-specific lump-sum transfers by considering a case where transfers are location- rather than individual-specific.[6] We also discuss how to apply our framework to dynamic and stochastic economies.

In Section 6, we compare our approach to common alternative measures of aggregate welfare used in the literature. We summarize these findings below.

i. **Real GDP**: Our measure generally differs from real GDP because, as is well-known, real GDP ignores amenity value. This means that it can be an inappropriate gauge for welfare (Roback, 1982).[7] For example, if an agent moves from a high- to a low-wage location/job, then real GDP can fall, although that agent may be better off once we take into account her location preferences. Indeed, it is simple to construct examples where a shock makes every agent better off, but causes real GDP to decline (e.g. if productivity rises in a low-wage location, then real output can fall because workers move there). We show that in perfectly competitive economies, our measure does coincide, to a first-order approximation, with measures of multi-factor productivity growth (or the Solow residual) computed using a quality-adjusted labor input. That is, our analysis provides a theoretical justification for using those statistics in efficient economies with heterogeneous tastes and discrete choice.

ii. **Sum of compensating variations**: As explained above, in contrast to the sum of compensating variations, our measure accounts for the fact that compensating transfers can alter equilibrium prices and wages. To illustrate this distinction, we show that a policymaker choosing place-based policies to maximize the sum of compensating variations will typically implement inefficient policies, taxing some locations to redistribute to others, even when the initial equilibrium is Pareto efficient.

iii. **Average utility**: The average utility measure does not coincide with our measure, even to first order. For example, in perfectly competitive economies, average utility does not obey Hulten's theorem (as pointed out by Donald et al., 2026). We show that the extent of these deviations from Hulten's theorem depends on which arbitrary but equally valid monotone transformations of individual utility functions are used. We use a version of the Allen and Arkolakis (2014) model with hetero-

[6] See also Schulz et al. (2023) who consider a mechanism design problem where compensations must be incentive compatible under imperfect information.

[7] Real GDP is nevertheless popular in the literature, partly because it is consistent with objects in national income accounting. Some recent examples that report real output include Hsieh et al. (2019), Lamadon et al. (2022), and Bagga et al. (2025) in models with occupation choice, and Bryan and Morten (2019) and Heise and Porzio (2022) in spatial migration models.

geneous tastes to quantitatively illustrate the difference between our approach and average utility. We show that these arbitrary assumptions also matter for policy prescriptions: a policymaker maximizing average utility can make extensive use of Pareto-inefficient place-based policies, even in Pareto efficient economies, and the optimal policy can be highly sensitive to how ordinal preferences are mapped into cardinal utils.

2 Environment and Equilibrium

In this section we define the environment and equilibrium. Households choose among locations or occupations that determine their earnings and consumption opportunities. Producers use primary factors and intermediate goods under constant returns and perfect competition. The general environment is fairly abstract, so readers looking for a concrete example are encouraged to focus on the running example introduced in Example 1. Throughout the rest of the paper, we return to this same example to illustrate the general results. Market imperfections and externalities are introduced in Section 5.

Commodity space. The commodity space contains a set C of consumption bundles, a set N of intermediate goods, and a set R of primary factors. Every commodity is indexed by a comprehensive set of characteristics that define its uniqueness including, for example, region (e.g. haircuts in LA and New York can be treated as different commodities).[8,9]

Primary factors are supplied by households and can be arbitrarily finely differentiated within R. For example, R can distinguish factors by region, occupation, skill-level, and type (e.g. high-skilled accountants in LA and commercial real estate in New York are different factors). Agents differ in their ability to provide these factor services. For example, some agents, called low-skilled workers, can only supply low-skilled labor, or some agents, called landowners in region x, can only supply land to firms in that region. Each agent h makes a discrete choice $l_h \in R$ that we refer to as the *location* of h, with the understanding that it may index location in space as well as occupation, industry, and so on.[10]

[8] In principle, this could include time and state of nature. That is, following Arrow-Debreu, our theorems and propositions can be applied to dynamic and stochastic economies. However, for concreteness, in all of our examples, we focus on static deterministic economies.

[9] Similarly, if there are trade costs, then traded goods are indexed by both origin and destination. For example, oranges produced in California and consumed in New York are different from oranges produced in California and consumed in Illinois.

[10] Because the set of factors, R, is kept abstract, this setup is quite general. For example, to capture commuting, where agents consume and work in different places (as in Monte et al., 2018), we let r index

For simplicity, the set of consumption bundles C has the same size as the set of primary factors R, with a one-to-one correspondence between consumption bundles and locations. Thus, for each $r \in R$, there is a primary factor ("labor in location r") and an associated consumption bundle $c(r) \in C$ ("consumption in location r"). By allowing consumption bundles to differ by location, we allow for the possibility that certain commodities are nontradeable or that there are trade costs (e.g. barbers in Los Angeles consume a different bundle of goods from waiters in New York).

Households' problem. There is a unit mass of agents, indexed by $h \in H$. Each agent chooses a location $l_h \in R$ and consumes a scalar quantity c_h of the homothetic composite associated with that location. Households choosing the same location consume the same composite, but their preferences over its quantity and the location itself may differ arbitrarily.

Agent h has ordinal preferences $\succeq_h$ over the location choice l_h and the quantity of the consumption bundle consumed. These preferences are represented by a utility function $u_h(c_h, l_h)$ that is continuous and strictly increasing without bound in consumption for each location, and is only pinned down up to a strictly increasing transformation. Each agent maximizes utility choosing consumption c_h and location l_h

$$\max_{c_h, l_h} u_h(c_h, l_h) \tag{1}$$

subject to a budget constraint that requires consumption expenditures be equal to factor income plus transfers:

$$p^c_{l_h} c_h = Z a_{h l_h} w_{l_h} + T_h,$$

where p^c_r denotes the price of the consumption bundle in location r, a_{hr} is agent h's efficiency units of factor type r, w_r is the wage per efficiency unit of factor r, and T_h is a household-specific lump-sum transfer independent of the household's location choice. We anticipate that firms earn zero profits in equilibrium. The scalar Z is an aggregate factor-augmenting productivity shifter that uniformly scales the efficiency of factors in every location for all agents.[11] We use Z to define our measure of aggregate welfare later.

Denote the utility-maximizing location choice of agent h by l_h. We refer to the quantity $a_{hr}\mathbf{1}[l_h = r]$ as labor supplied by h, with the understanding that this is just a shorthand and factors could in principle include non-labor inputs like land. Given this shorthand,

pairs of locations instead. In a dynamic or stochastic model, we would let r index a state-contingent plan of locations.

[11] Equivalently, we could model Z as a shifter that scales factor-augmenting productivity of all producers.

efficiency units of labor supplied in location r (not including the aggregate shifter Z) are

$$L_r = \int a_{hr} \mathbf{1}[l_h = r] dh. \tag{2}$$

Efficiency units of labor supplied and available for production in location r are ZL_r. Final consumption spending in location r is

$$E_r = Z w_r L_r + \int T_h \mathbf{1}[l_h = r] dh. \tag{3}$$

Producers' problem. The set of produced commodities is the union of consumption goods, C, and intermediate goods, N, denoted by $C + N$. Each commodity i is produced by a perfectly competitive representative firm that maximizes its profit:

$$\max_{y_i, x_{ij}, l_{ir}} p_i y_i - \sum_{j \in N} p_j x_{ij} - \sum_{r \in R} w_r l_{ir}, \tag{4}$$

subject to a constant-returns technology

$$y_i = z_i F_i \left(\{x_{ij}\}_{j \in N}, \{l_{ir}\}_{r \in R} \right), \tag{5}$$

where z_i is a Hicks-neutral productivity shifter for producer i, F_i is a constant returns production function, x_{ij} is producer i's use of intermediate good j, and l_{ir} is its use of efficiency units of primary factor r. The price of the consumption bundle in location $r \in R$ is $p_r^c = p_{c(r)}$.

Market clearing constraints. The resource constraint requires that consumption and intermediate input usage be equal to production. Recall that commodities are split into pure consumption goods ($i \in C$) and pure intermediates ($i \in N$). Hence, if i is a consumption bundle, then i is not used as an intermediate ($x_{ji} = 0$ for each j). On the other hand, if $i \in N$, then i is not directly consumed by any household. The market clearing constraints are therefore:

$$E_r = p_r^c \int c_h \mathbf{1}[l_h = r] dh = p_r^c y_{c(r)} \text{ if } r \in R \qquad \text{and} \qquad p_i \sum_{j \in C+N} x_{ji} = p_i y_i \text{ if } i \in N. \tag{6}$$

The resource constraint for factor $r \in R$ requires that total factor demand from producers is equal to total factor supply from households:

$$\sum_{i \in C+N} l_{ir} = ZL_r. \tag{7}$$

Market clearing constraints imply that lump-sum transfers add up to zero:

$$\int T_h dh = 0. \tag{8}$$

We can now define a competitive equilibrium.

Definition 1 (Competitive Equilibrium)**.** For given productivity shifters z and Z, a competitive equilibrium with transfers is a collection of consumptions, c_h, location choices, l_h, transfers, $\{T_h\}$, outputs, y_i, input choices, $\{x_{ij}, l_{ir}\}$, prices, p_i, and wages, w_r, such that each agent chooses consumption and location to maximize utility subject to their budget constraint, producers maximize profits subject to technology taking prices as given, transfers add up to zero, and all markets clear. That is, equations (1)-(8) must be satisfied.

Example 1 (One-Good Economy). There is a continuum of agents with heterogeneous preferences represented by utility functions $u_h(c_h, l_h)$. For example, the utility function could take the form:

$$u_h(c_h, l_h) = f_h(g(c_h) + \epsilon_{hl_h}),$$

where g is increasing, f_h is any arbitrary monotone increasing function, and ϵ_{hl_h} is a parameter capturing h's tastes for choosing location l_h. There is a single freely-traded consumption good produced, which we take as the numeraire. Agents have the same skills, $a_{hr} = 1$ for every h and r, and there are no initial transfers. Each agent's budget constraint is

$$c_h = Zw_{l_h}.$$

There is one producer in each location with a linear production technology:

$$y_r = Zz_r \int \mathbf{1}[l_h = r] dh = Zz_r L_r,$$

where z_r is the productivity and L_r is the supply of labor in location r. Consumption goods produced in each location are perfectly substitutable and freely traded, so the ag-

gregate resource constraint is therefore total consumption equals total production:

$$\int c_h dh = Z \sum_r z_r L_r.$$

In the competitive equilibrium, the wage is equal to local productivity $w_r = z_r$. Households consume $c_h = Z z_{l_h}$ and choose

$$l_h \in \arg\max_{r \in R} u_h(Z z_r, r).$$

Thus, households trade off consumption, determined by local productivity, against their preferences over locations.

---

3 Defining and Characterizing Aggregate Welfare

In this section, we define our measure of aggregate welfare and then characterize it. Consider an initial *status quo* equilibrium allocation, corresponding to the allocation observed in the data. Let $\boldsymbol{z}^0$ be the vector of producer-specific productivity shifters in the status quo. Similarly, denote all other variables x at the status quo by x^0. Normalize the aggregate factor-augmenting productivity shifter to equal one in the status quo, $Z^0 = 1$. We are interested in measuring the aggregate economic surplus created by a change in the vector of productivity parameters from $\boldsymbol{z}^0$ to $\boldsymbol{z}$.

3.1 Definition of TFP-Equivalent Aggregate Surplus

Define the set of feasible allocations, given productivities $\boldsymbol{z}$ and Z, to be

$$\mathcal{C}(\boldsymbol{z}, Z) \equiv \{\{c_h, l_h\}_{h \in H} \text{ supported via competitive equilibrium given } \boldsymbol{z}, Z, \text{ and some transfers}\}.$$

Feasible here means allocations that can be implemented as competitive equilibrium allocations given some transfers (not just technologically feasible).

Example 2 (One-Good Economy's Feasible Set). Consider the economy described in Example 1: there is a single freely-traded consumption good produced linearly from labor

in each location. The feasible consumption set for this economy is:

$$\mathcal{C}(\boldsymbol{z}, Z) = \Bigg\{ \{c_h, l_h\}_{h\in H} : \text{there exist } \{T_h\}_{h\in H} \text{ such that } \int T_h \, dh = 0, \\ c_h = Z z_{l_h} + T_h, \quad l_h \in \arg\max_{r\in R} u_h \left(Z z_r + T_h, r\right) \Bigg\}.$$

The first condition is that transfers must be budget-balanced, the second gives households' budget constraints, and the last states that households choose their locations to maximize utility.

We measure changes in aggregate welfare in response to a change in productivities $\boldsymbol{z}$ using equivalent changes in factor-augmenting productivity Z.

Definition 2 (TFP-Equivalent Aggregate Surplus)**.** The aggregate surplus generated by the change, in TFP-equivalent terms, is defined to be

$$A = \max \left\{ Z^{-1} : \{c_h, l_h\}_{h\in H} \in \mathcal{C}(\boldsymbol{z}, Z) \text{ and } (c_h, l_h) \succeq_h (c_h^0, l_h^0) \text{ for every } h \right\}, \tag{9}$$

where c_h^0 and l_h^0 are the consumption and location of agent h in the status quo.

In words, to measure the change in aggregate welfare, we choose the smallest common factor-productivity multiplier Z that permits an equilibrium with budget-balanced transfers in which every household is at least as well off as in the status quo.[12] If A is greater than one, say $A = 1.01$, then this means that we can discard roughly 1% of every factor endowment (more precisely, $1 - 1.01^{-1}$) and still keep every household indifferent to the initial equilibrium. Equivalently, because Z uniformly scales the production possibility set, every household can be kept indifferent to the status quo with roughly 1% of every good left-over. In other words, this represents an aggregate surplus gain equivalent to a 1% increase in TFP. Note that this quantity is invariant to monotone transformations of utility functions, and is distribution neutral in that it does not take a stance on who should receive the surplus. When A is less than one, instead of measuring the surplus, A measures the shortfall in factor endowments needed to make everyone whole.

[12]This definition is similar to the one used by Baqaee and Burstein (2023) in the context of continuous choice economies. It is not exactly the definition in Debreu (1951) because we restrict the set of feasible allocations to be the ones that can be reached using individual-specific lump-sum transfers as opposed to any technologically feasible allocation. In this economy, the second welfare theorem need not hold, so these two definitions need not be equivalent. In Section 5, we discuss how to further restrict the set of feasible allocations to be the ones that can be supported given market imperfections or place-based, rather than individual-specific, transfers.

Throughout, we assume that at the solution to (9), every agent can be made exactly indifferent to the status quo. This rules out pathological cases where agents are completely disconnected in goods and factor markets.[13]

Before proceeding with the analysis, we collect some remarks on how this measure can be extended.

Remark (Changes in amenities). Our analysis focuses on the effect of productivity shocks. However, the same definition of aggregate welfare and many of our results can be extended to cover changes in other parameters as well. For example, we can allow for changes in location characteristics that affect amenity values (e.g. changes in the weather in a region or in the health risk associated with particular occupations). For example, if the amenity value of choosing location r can be written as $u_h(B_r c_r, r)$, where B_r is a location-specific quality shifter, then changes in B_r have the same effect on aggregate surplus A as changes in the productivity of the consumption good in that location, $z_{c(r)}$.

Remark (Welfare for a subset of agents). Definition (9) can also be applied to measure collective welfare for a subset of agents, such as agents of a given gender, race, or age. In that case, we contract the endowments of agents in that subset subject to indifference conditions and balanced-budget transfers within the subset. If the subset consists of a single infinitesimal agent h, then the TFP-equivalent surplus collapses to that agent's compensating variation,[14] denoted by A_h, and defined by

$$\max_r u_h \left(\frac{a_{hr} w_r}{A_h p_r^c}, r \right) = u_h^0. \tag{10}$$

The value of A_h depends only on household h's ordinal preferences $\succeq_h$ and is invariant to strictly increasing transformations of utility.

3.2 Characterization of Aggregate Welfare A

To solve for A, we break the problem into two parts. First, we compute where households choose to go and how much they spend once they are compensated, holding relative prices constant. As in classical consumer theory, this gives us the compensated supply and demand functions, which map wages and prices into labor supplied and spending

[13] This happens if, for example, households are immobile across locations and each location is in autarky. In such a case, making households exactly indifferent to the status quo can become impossible because transfers between locations are not possible. In these cases the measure is itself still well-defined, but some of our theorems do not go through.

[14] Technically, this is a compensating variation in ratios—the proportional reduction in agent h's productivity—rather than in differences. See Appendix D.2 for a comparison of compensating variations in ratios and in differences.

in each location, holding each agent on a given indifference curve.[15] Second, given those compensated supply and demand choices, we solve for relative prices that can clear markets. The value of A is then pinned down by a fixed point where compensated supply and demand are consistent with market clearing. In the final part of this section, we show how one can recover the aggregate compensated supply and demand functions, needed to compute A, from aggregate uncompensated labor supply functions, which are much easier to estimate.

3.2.1 Expenditure Function and Compensated Choices

Begin by defining the expenditure function of each consumer.

Definition 3 (Expenditure function)**.** Define $T_h(\boldsymbol{w}, \boldsymbol{p}^c)$ to be the (net) expenditure function for agent h given wages $\boldsymbol{w}$ and consumption prices $\boldsymbol{p}^c$:

$$T_h(\boldsymbol{w}, \boldsymbol{p}^c) = \min_{l_h, c_h} \left\{ p^c_{l_h} c_h - a_{hl_h} w_{l_h} : (c_h, l_h) \succeq_h (c^0_h, l^0_h) \right\}.$$

We call the location choice $l^{\text{comp}}_h(\boldsymbol{w}, \boldsymbol{p}^c)$ associated with this optimization problem *the compensated choice* of agent h.

In words, $T_h(\boldsymbol{w}, \boldsymbol{p}^c)$ is the minimum lump-sum transfer agent h requires to be made at least indifferent to the status quo allocation, given prices and wages. The location choice that the agent makes, given that transfer, is what we call the *compensated* choice. This definition parallels the classical definition of expenditure functions and compensated demand in consumer theory.[16]

The following proposition shows how to compute the compensated choices and the expenditure function agent-by-agent.

Proposition 1 (Compensated Choices and Expenditure Function)**.** *The compensated choice* $l^{comp}_h(\boldsymbol{w}, \boldsymbol{p}^c)$ *satisfies*

$$l^{comp}_h(\boldsymbol{w}, \boldsymbol{p}^c) \in \arg\max_{r \in R} \left\{ a_{hr} w_r - p^c_r \bar{c}_{hr} \right\},$$

[15] In Appendix A.2, we consider a special case with homogeneous agents. In this case, lump-sum transfers are not needed to evaluate the compensated supply and demand functions, since all agents are affected symmetrically by shocks. In this case, if households only care about consumption and do not directly value location choices, then A is just the growth in real consumption in any non-empty location (see Corollary 6 in Appendix A.2).

[16] See also Small and Rosen (1981) for a related analysis. Whereas they consider discrete consumption choices, given fixed income, here the level of income depends also on the discrete choice (via the wage).

where for each $r \in R$, $\bar{c}_{hr}$ is the solution to $u_h(\bar{c}_{hr}, r) = u_h(c_h^0, l_h^0)$. The expenditure function, or transfer needed to ensure indifference to u_h^0, is

$$T_h(\boldsymbol{w}, \boldsymbol{p}^c) = p^c_{l_h^{comp}} \bar{c}_{h l_h^{comp}} - a_{h l_h^{comp}} w_{l_h^{comp}}.$$

In words, $(a_{hr} w_r - p_r^c \bar{c}_{hr})$ is the surplus, in dollars, agent h receives from being sent to location r, where $\bar{c}_{hr}$ is the consumption agent h must have in location r to be indifferent to the status quo. The compensated choice maximizes this surplus, and the expenditure function is equal to the negative of the surplus. Proposition 1 gives a straightforward way to calculate the compensated choice and the compensating transfer for each agent given any vector of consumption prices $\boldsymbol{p}^c$ and wages $\boldsymbol{w}$.

We now aggregate agent-level expenditure functions and compensated choices.

Definition 4 (Aggregate Expenditure & Choices)**.** The aggregate expenditure function is

$$\overline{T}(\boldsymbol{w}, \boldsymbol{p}^c) = \int T_h(\boldsymbol{w}, \boldsymbol{p}^c) dh.$$

The aggregate compensated labor supply function is:

$$L_r^{\text{comp}}(\boldsymbol{w}, \boldsymbol{p}^c) = \int a_{hr} \mathbf{1}[l_h^{\text{comp}}(\boldsymbol{w}, \boldsymbol{p}^c) = r] dh.$$

Aggregate compensated demand in location r is:

$$E_r^{\text{comp}}(\boldsymbol{w}, \boldsymbol{p}^c) = w_r L_r^{\text{comp}}(\boldsymbol{w}, \boldsymbol{p}^c) + \int T_h(\boldsymbol{w}, \boldsymbol{p}^c) \mathbf{1}[l_h^{\text{comp}}(\boldsymbol{w}, \boldsymbol{p}^c) = r] dh,$$

where $l_h^{\text{comp}}(\boldsymbol{w}, \boldsymbol{p}^c)$ and $T_h(\boldsymbol{w}, \boldsymbol{p}^c)$ are given by Proposition 1.

The following proposition relates aggregate expenditures to the aggregate compensated supply and demand functions by way of an aggregate version of Shephard's lemma. Throughout, we assume that, at the wage and price vectors under consideration, compensated location choices are unique for almost every household.[17]

Proposition 2 (Aggregate Shephard's Lemma)**.** *The aggregate compensated labor supply function satisfies*

$$\frac{\partial \overline{T}(\boldsymbol{w}, \boldsymbol{p}^c)}{\partial w_r} = -L_r^{comp}(\boldsymbol{w}, \boldsymbol{p}^c),$$

[17] This ensures that the aggregate expenditure function is differentiable.

and the aggregate compensated demand satisfies

$$\frac{\partial \overline{T}(\boldsymbol{w}, \boldsymbol{p}^c)}{\partial \log p_r^c} = E_r^{comp}(\boldsymbol{w}, \boldsymbol{p}^c).$$

Moreover, since the aggregate expenditure function is homogeneous of degree one in wages and prices,

$$\overline{T}(\boldsymbol{w}, \boldsymbol{p}^c) = \sum_r \left(E_r^{comp}(\boldsymbol{w}, \boldsymbol{p}^c) - w_r L_r^{comp}(\boldsymbol{w}, \boldsymbol{p}^c) \right).$$

In words, the derivatives of the aggregate expenditure function reveal compensated supply and demand functions, and since expenditures are homogeneous of degree one, these derivatives can also be used to reconstruct aggregate expenditures. The next result shows that the aggregate compensated supply and demand functions, or equivalently the aggregate expenditure function, can be used to calculate A.

3.2.2 Calculating Aggregate Welfare Using Compensated Supply and Demand

The next result uses the compensated supply and demand functions to solve for A.

Theorem 1 (Aggregate Welfare using Compensated Equilibrium)**.** *TFP-equivalent aggregate surplus, A, solves*

$$\sum_r E_r^{comp}(\boldsymbol{w}/A, \boldsymbol{p}^c) = \sum_r (w_r/A) L_r^{comp}(\boldsymbol{w}/A, \boldsymbol{p}^c), \tag{11}$$

where prices clear the modified market clearing conditions

$$E_r^{comp}(\boldsymbol{w}/A, \boldsymbol{p}^c) = p_r^c y_{c(r)} \text{ if } r \in R \qquad \text{and} \qquad p_i \sum_{j \in C+N} x_{ji} = p_i y_i \text{ if } i \in N. \tag{12}$$

and wages clear the modified labor market clearing conditions

$$w_r \left(\sum_{i \in C+N} l_{ir} \right) = \frac{w_r}{A} L_r^{comp}(\boldsymbol{w}/A, \boldsymbol{p}^c). \tag{13}$$

In these market clearing conditions, the firm side consisting of intermediate input demand, factor demand, and production technologies is the same as in the decentralized equilibrium.

The wages, prices, and quantities that satisfy these conditions are not the same as the ones in the decentralized equilibrium. We call the collection of prices, wages, and quantities that satisfy the conditions in Theorem 1 the *compensated equilibrium*.

Intuitively, in the compensated equilibrium, we replace household supply and demand functions by their compensated variants in (12) and (13). We then adjust total factor endowments, via A, to ensure that total spending equals total earnings (11). Equivalently, we adjust factor endowments to ensure that lump-sum transfers across households add up to zero. Hence, conditional on knowing the aggregate compensated supply and demand functions $\boldsymbol{L}^{\text{comp}}$ and $\boldsymbol{E}^{\text{comp}}$, solving for a compensated equilibrium is only slightly more difficult than solving for a competitive equilibrium — there is only one additional equation, (11), and one additional unknown, A.

Note that (11) is equivalent to saying that $1/A$ is the level of TFP that sets net aggregate expenditures (spending in excess of income) to zero:

$$\overline{T}(\boldsymbol{w}/A, \boldsymbol{p}^c) = 0.$$

The wages and consumption prices in this expression are the ones in a compensated equilibrium where we use compensated aggregate labor supply and demand in the equilibrium conditions. Compensated supply and demand are evaluated at $\boldsymbol{w}/A$ because those are the wages households face per efficiency unit of labor that they supply. Since the aggregate compensated supply and demand functions are themselves partial derivatives of the aggregate expenditure function (Proposition 2), the aggregate expenditure function carries all the information from the household-side of the model needed to compute A.

We illustrate Theorem 1 using a simple example.

Example 3 (Compensated Equilibrium in One-Good Economy). Consider the economy described in Examples 1 and 2: there is a single freely-traded consumption good produced linearly from labor in region r with productivity z_r. Applying Theorem 1 is simple in this economy since the real wage (per efficiency unit) in each location in the compensated equilibrium is equal to exogenous parameters: z_r. Hence, (11) can be rearranged as

$$A = \frac{\sum_r z_r L_r^{\text{comp}}(\boldsymbol{z}/A)}{\sum_r E_r^{\text{comp}}(\boldsymbol{z}/A)} = \frac{\sum_r \int z_r \mathbf{1}[l_h^{\text{comp}} = r]dh}{\sum_r \int \bar{c}_{hr} \mathbf{1}[l_h^{\text{comp}} = r]dh}, \tag{14}$$

which is the ratio of production (given the technologies at $\boldsymbol{z}$ but compensated locations) and the amount of consumption necessary to keep everyone indifferent at those locations. By Proposition 1, the compensated location choice of agent h is

$$l_h^{\text{comp}} = \arg\max_{r \in R} \left\{ z_r/A - \bar{c}_{hr} \right\}.$$

This is a simple numerical problem. For example, we first solve for l_h^{comp} agent-by-agent given a conjectured value of A, and then use (14) to update the value of A until we reach a fixed point.

The requirement that compensating transfers sum to zero also gives A an interpretation in terms of producer and consumer surplus. Households supply labor and consume goods, so changes in effective wages and consumption prices generate producer- and consumer-surplus gains, respectively. The following corollary shows that scaling factor productivity by $1/A$ makes these gains sum to zero in the compensated equilibrium.

Corollary 1 (Consumer and producer surplus)**.** *The change in aggregate welfare, A, is the adjustment in factor productivity that makes the sum of producer- and consumer-surplus for households equal to zero in the compensated equilibrium. Formally, let $\boldsymbol{z}(t)$, for $t \in [0,1]$, trace a path from $\boldsymbol{z}^0$ to $\boldsymbol{z}$, and suppose the associated compensated equilibrium varies smoothly along this path. Then*

$$\underbrace{\int_0^1 \sum_r L_r^{\text{comp}}(t) \frac{d}{dt}\left(\frac{w_r(t)}{A(t)}\right) dt}_{\textit{Producer-surplus gains}} + \underbrace{\left[-\int_0^1 \sum_r E_r^{\text{comp}}(t) \frac{d \log p_r^c(t)}{dt} dt\right]}_{\textit{Consumer-surplus gains}} = 0, \tag{15}$$

where $w_r(t)$ and $p_r^c(t)$ are wages and prices in the compensated equilibrium.

The first integral accumulates changes in the area above compensated factor supply and below effective wages. The second accumulates changes in the area below compensated consumption demand and above consumption prices. The role of A is to adjust productive capacity until the combined producer- and consumer-surplus gain is zero. If the shock generates a surplus, factor endowments can be contracted by $1/A < 1$ until that surplus is exhausted. If the shock generates a loss, factor endowments must instead expand.

Proof. By Proposition 2, the following is an identity:

$$d\overline{T} = -\sum_r L_r^{\text{comp}}\, d(w_r/A) + \sum_r \frac{E_r^{\text{comp}}}{p_r^c}\, dp_r^c.$$

Integrating and using $\overline{T}(\boldsymbol{w}(t)/A(t), \boldsymbol{p}^c(t)) = 0$ at both endpoints, gives (15). □

Now, we consider a special case where aggregate surplus can be computed from the uncompensated aggregate supply function directly.

Proposition 3 (Aggregate Welfare from Uncompensated Labor Supply)**.** *Suppose all households consume the same good and preferences can be represented by $u_h(c, r) = f_h(c + \epsilon_{hr})$, where f_h is strictly increasing. Normalize the consumption price to one, and let $L_r(\boldsymbol{w})$ denote uncompensated aggregate labor supply as a function of effective real wages.*

The compensated equilibrium has the same wages, production, and aggregate labor supplies as the decentralized equilibrium with aggregate factor productivity $1/A$. Along the path in Corollary 1, aggregate welfare satisfies

$$\int_0^1 \sum_r L_r \left(\frac{\boldsymbol{w}(t)}{A(t)} \right) \frac{d}{dt} \left(\frac{w_r(t)}{A(t)} \right) dt = 0. \tag{16}$$

Thus, A can be computed using only uncompensated labor supply and the production side of the economy.

Proposition 3 follows from two simplifications. First, the restriction that there be only one consumption good means that a lump-sum transfer adds the same amount to consumption in every location and, given the restriction on preferences, this leaves each household's location choice unchanged. Hence, the first simplification is that the compensated and uncompensated labor supply coincide. Second, because everyone consumes the same good, redistributing income across locations does not change the composition of aggregate demand or relative prices in equilibrium. We can therefore use the ordinary equilibrium conditions and determine A by using Corollary 1, which requires that the net surplus be zero. In this example, this simplifies to (16) using the uncompensated labor supply function.[18]

The following example illustrates Proposition 3.

Example 4 (Surplus in the One-Good Economy). Consider Example 1 again: there is one common and freely traded consumption good made linearly from labor in each location, which we take as the numeraire. In this economy, effective real wages are pinned down by productivity parameters only: $w_r/(Ap_r^c) = z_r/A$. Suppose that preferences take the

[18] In Appendix A.3 we show that under the assumptions of Proposition 3, A can also be computed using a utilitarian social welfare function $U(\boldsymbol{w}/\boldsymbol{p}^c) = \mathbb{E}[\max_{l_h} a_{hl_h} w_{l_h}/p^c_{l_h} + \epsilon_{h,l_h}]$ where $a_{hr} w_r/p_r^c$ is real consumption of household h when choosing location r. This specific utilitarian function U is called the "*Social Surplus Function*" in the discrete choice literature (see McFadden, 1981). The reason is that, by the Williams-Daly-Zachary theorem, the social surplus function is the integral of uncompensated labor supply. Under the assumptions of Proposition 3, the compensated and uncompensated labor supply coincide, so the social surplus function is also the integral of compensated supply. Hence, by a similar logic to (16), $U(\boldsymbol{w}/\boldsymbol{p}^c)$ can be used to determine A. However, if the assumptions of Proposition 3 do not hold, then the social surplus function cannot in general be used to compute A.

form in Proposition 3, so that we can use the uncompensated supply function in place of the compensated supply function. This implies that A is pinned down by:

$$\int_0^1 \sum_r L_r \left(\frac{\boldsymbol{z}(t)}{A(t)} \right) \frac{d}{dt} \left(\frac{z_r(t)}{A(t)} \right) dt = 0.$$

If we know the functional form for the uncompensated supply function, then the equation above is a single nonlinear equation that pins down A. For example, suppose preferences take the form $u_h(c_h, l_h) = f_h(c_h + \epsilon_{hl_h})$, where ϵ_{hr} are independently drawn from type-I extreme-value (i.e., Gumbel) distributions. Following McFadden (1973), the aggregate uncompensated labor supply in each location is

$$L_r = \frac{\exp(\theta Z w_r / p^c + B_r)}{\sum_{r'} \exp(\theta Z w_{r'} / p^c + B_{r'})}, \tag{17}$$

where θ and $\{B_r\}$ are parameters. With this functional form, we can evaluate the integral in closed-form as:

$$\log \sum_r \exp \left(\theta \frac{z_r}{A} + B_r \right) = \log \sum_r \exp \left(\theta z_r^0 + B_r \right). \tag{18}$$

This nonlinear equation pins down A.

3.3 From Uncompensated to Compensated Choices

Proposition 3 shows that under some strong assumptions, we can avoid carrying out the compensation calculations because transfers do not change household choices. Outside of this special case, Theorem 1 and Proposition 2 show that the aggregate expenditure function $\overline{T}(\boldsymbol{w}, \boldsymbol{p}^c)$ summarizes the household side of the model for the purposes of computing A. To evaluate the aggregate expenditure function, one can always use Definition 4 combined with Proposition 1. In other words: simulate a large population of agents, compute their individual expenditure functions and compensated choices, given wages and prices, then add up the results. We now show that, more generally, the aggregate net expenditure function can be recovered from uncompensated choices, without simulating individual households.

For simplicity, we focus on the case without skill-heterogeneity. This implies that agents make different choices only due to differences in tastes. If agents have different

skills, then the same ideas apply, but conditional on skill level.

Assumption 1 (Homogeneous Skills)**.** Suppose that skills do not vary across households: $a_{hr} = a_r$ for every $h \in H$ and $r \in R$. Without loss of generality, we normalize $a_r = 1$ for every $r \in R$.

Let $L_r(\boldsymbol{c})$ denote uncompensated labor supply in location r as a function of a vector of consumption across locations, $\boldsymbol{c} = (c_1, \ldots, c_R)$. Given wages, consumption prices, and a common lump-sum transfer T, consumption in location r is $c_r = (w_r + T)/p_r^c$ (under the assumption $a_{hr} = 1$). The following proposition shows that uncompensated labor supply pins down the aggregate expenditure function $\overline{T}(\boldsymbol{w}, \boldsymbol{p}^c)$ and, hence, the aggregate compensated supply and demand functions.

Proposition 4 (Aggregate Expenditure Function from Uncompensated Labor Supply)**.** *Suppose that lump-sum transfers in the status quo are zero and that Assumption 1 holds. The aggregate expenditure function is given by*

$$\overline{T}(\boldsymbol{w}, \boldsymbol{p}^c) = T^{min}(\boldsymbol{w}, \boldsymbol{p}^c) + \int_{T^{min}(\boldsymbol{w}, \boldsymbol{p}^c)}^{T^{max}(\boldsymbol{w}, \boldsymbol{p}^c)} \sum_r L_r(\widetilde{\boldsymbol{c}}(x)) \, \mathbf{1}\left[c_r^0 > \frac{w_r + x}{p_r^c}\right] dx, \tag{19}$$

where

$$\widetilde{c}_r(x) \equiv \max\left\{c_r^0, \frac{w_r + x}{p_r^c}\right\},$$

and

$$T^{min}(\boldsymbol{w}, \boldsymbol{p}^c) \equiv \min_r \left\{p_r^c c_r^0 - w_r\right\}, \qquad T^{max}(\boldsymbol{w}, \boldsymbol{p}^c) \equiv \max_r \left\{p_r^c c_r^0 - w_r\right\}.$$

Equation (19) can be understood as follows. For a given transfer x, the number $\widetilde{c}_r(x)$ is the larger of status-quo consumption in location r, c_r^0, and counterfactual consumption after transfer x. Households who, facing the vector $\widetilde{\boldsymbol{c}}(x)$, choose a location where the first term is larger are those for whom x is not sufficient compensation to make them indifferent to the status quo. Therefore, the sum inside the integral (19) is the share of households that require a compensating transfer greater than x. This step uses an insight from Bhattacharya (2015, 2018). The result then follows from the tail-sum formula for expectations which relates the expectation of a random variable, in this case $\int T_h dh$, to its countercumulative distribution function, in this case the share of households for whom T_h is greater than x. For more details, see the proof in Appendix A.1.

Proposition 4 is useful because it shows that to calculate aggregate welfare, A, the aggregate uncompensated labor supply function L_r is a sufficient statistic for the household side of the model. In particular, one does not need to specify individual-level preferences

or estimate micro-level labor supply. Instead, one can directly estimate aggregate uncompensated labor supply from aggregate (that is, market-level) variation in employment in response to exogenous changes in real wages.

It is straightforward to extend Proposition 4 to heterogeneous skills by applying the result within each skill type and then aggregating across types.

4 Hulten's Theorem and Beyond

In this section, we characterize aggregate welfare exactly as a cumulative sales-weighted sum of productivity changes in a compensated equilibrium. We then derive first- and second-order approximations and identify the expenditure shares and substitution elasticities needed to compute them.

Theorem 2 (Nonlinear Hulten's Theorem)**.** *Let $\boldsymbol{z}(t)$, for $t \in [0,1]$, trace any smooth path from $\boldsymbol{z}^0$ to $\boldsymbol{z}^1$, and suppose the associated compensated equilibrium varies smoothly along this path. Let $\lambda_i^{comp}(t)$ denote the sales of producer i relative to GDP in the compensated equilibrium at t. The TFP-equivalent welfare change is*

$$\log A = \int_0^1 \sum_i \lambda_i^{comp}(s) \frac{d \log z_i}{ds} ds. \tag{20}$$

In words, the TFP-equivalent welfare change is the chained Domar-weighted sum of productivity shocks in the compensated equilibrium.[19] At the status quo, $t = 0$, compensated Domar weights coincide with decentralized Domar weights $\lambda_i^{\text{comp}}(0) = \lambda_i^0$.[20] This allows us to establish the following useful corollary of Theorem 2.

Corollary 2 (First-Order Hulten's Theorem)**.** *To a first-order approximation in $\Delta \log \boldsymbol{z} = \log z - \log z^0$, the TFP-equivalent welfare change satisfies*

$$\log A \approx \sum_i \lambda_i^0 \Delta \log z_i,$$

where λ_i^0 is the observed sales of i divided by GDP in the decentralized equilibrium at the status quo.

[19] In Appendix B.4, we show that Theorem 2 can be used to compute A as a solution to a system of ordinary differential equations. Solving these differential equations requires knowledge of expenditure shares (including the input-output table) at the status quo, as well as substitution elasticities (those in production, as well as the compensated supply and demand systems).

[20] This is because the decentralized equilibrium is Pareto efficient (the first welfare theorem holds), and hence, $A(0) = 1$ and the decentralized equilibrium supports the compensated allocation with no additional transfers needed.

Thus, to first order, aggregate welfare growth equals the sum of producers' productivity growth weighted by their initial sales relative to GDP. This coincides with multi-factor productivity growth measured using quality-adjusted labor inputs. We return to this relationship between A and multifactor productivity growth in Section 6.1 where we compare A to real GDP.

Corollary 2 is a generalization of Hulten's theorem to economies with discrete choice. Using Theorem 2 beyond first-order approximations requires knowledge of sales shares, $\lambda_i^{\text{comp}}(t)$, in the compensated equilibrium away from status quo $t > 0$. This is easy to do, using standard methods, similar to Baqaee and Farhi (2024), if we know the elasticities of compensated supply and demand. Under some assumptions, we can guarantee that λ_i^{comp} is constant in the compensated equilibrium, which allows us to turn Corollary 2 into an exact nonlinear expression.

Corollary 3 (Cobb-Douglas Economies with Common Consumption)**.** *Suppose that all production functions are Cobb-Douglas, that there is a single freely traded consumption good, produced using a Cobb-Douglas technology. Then*

$$\log A = \sum_i \lambda_i^0 [\log z_i - \log z_i^0].$$

Conditional on initial sales shares, no additional information about preferences, skills, or the input-output structure is needed.

Cobb-Douglas technologies keep producers' input expenditure shares fixed. With a common consumption good, redistribution across households does not change the composition of aggregate final demand. Consequently, Domar weights remain constant along the compensated equilibrium path and can be taken outside the integral in (20).

In general, however, the compensated sales shares do respond to changes in productivity shifters in the compensated equilibrium. The next result quantifies these nonlinear effects.

Corollary 4 (Second-Order Approximation)**.** *To a second-order approximation in productivity shocks, the TFP-equivalent welfare change satisfies*

$$\Delta \log A \approx \sum_i \lambda_i^0 \Delta \log z_i + \frac{1}{2} \sum_i \Delta \lambda_i^{comp} \Delta \log z_i,$$

where $\Delta \lambda_i^{comp}$ is the first-order change in the i's sales-to-GDP ratio due to the productivity shocks in the compensated equilibrium.

The intuition is very similar to the formulas derived by Baqaee and Farhi (2019) for an economy without discrete choice. If sales shares in the compensated equilibrium rise in response to positive productivity shocks, then aggregate welfare rises more quickly than its first-order approximation would suggest. If sales shares fall, then the reverse happens. In Appendix B.3, we show how to solve out for the change in compensated sales shares, $\Delta\lambda_i^{\text{comp}}$, in terms of the initial expenditure shares (the input-output table) and uncompensated price and wage elasticities of supply and demand.

We illustrate Corollary 4 with a simple example.

Example 5 (Second-Order Approximation in One-Good Economy). Consider again the one-good economy described in Example 1. Applying Corollary 4 to the one-good economy we get that, to a second-order,

$$\Delta \log A \approx \mathbb{E}_\lambda[\Delta \log z] + \frac{1}{2}\left[Var_\lambda\left(\Delta \log z\right) + Cov_\lambda(\Delta \log L^{\text{comp}}, \Delta \log z)\right], \tag{21}$$

where the variance and covariance operators use as weights the initial sales shares $\lambda_r^0 = z_r^0 L_r^0 / \sum_s z_s^0 L_s^0$, which sum to one. We also use the fact that $d \log \lambda_r^{\text{comp}} = (d \log(z_r L_r^{\text{comp}}) - \mathbb{E}_\lambda[d \log(z L^{\text{comp}})])$. In words, (21) states that the nonlinear change in aggregate welfare depends on the variance of the shocks and the covariance of compensated migrations with the shocks. The variance term reflects the fact that, even if households are not mobile, an increase in productivity in one location raises that location's sales share, which magnifies the value of positive productivity growth in that location relative to the loglinear approximation. The covariance reflects the fact that if, in the compensated equilibrium, households move to locations experiencing more rapid productivity growth, then this further boosts the compensated sales shares of producers in that location, and raises aggregate welfare.

To illustrate further, let us consider a specific example. Suppose that $u_h(c_h, r) = f_h(B_r c_h \varepsilon_{hr})$, where f_h is strictly increasing, ε_{hr} is i.i.d. unit-scale Fréchet distributed with shape parameter θ, and the parameters $\{B_r\}$ control the desirability of living in each location. Following Redding and Rossi-Hansberg (2017), the share of households that choose location r in the decentralized equilibrium takes an isoelastic form

$$L_r(\boldsymbol{z}) = \frac{(B_r z_r)^\theta}{\sum_{r'}(B_{r'} z_{r'})^\theta}. \tag{22}$$

For simplicity, suppose that initial productivities, z_r^0, are the same in every location. As

mentioned above, Appendix B.3 (specifically Proposition 14) provides the elasticities of compensated supply in terms of the elasticities of the uncompensated supply. In this example, this yields

$$\Delta \log L_r^{\text{comp}} = \theta \left(\Delta \log z_r - \mathbb{E}_\lambda \left[\Delta \log z \right] \right).$$

Substituting this into the approximation for $\Delta \log A$ in (21) yields:

$$\Delta \log A \approx \mathbb{E}_\lambda [\Delta \log z] + \frac{1}{2}(1+\theta) Var_\lambda \left(\Delta \log z \right).$$

Holding initial sales shares fixed, a uniform productivity change has no second-order correction. For uneven changes, the correction is positive and increases with the mobility elasticity θ. The term 1 captures changing sales shares even without relocation; the term θ captures the additional contribution of labor moving toward locations with larger productivity gains.

---

5 Extensions

In this section, we provide some extensions to the basic framework. We show how to extend the framework to allow for externalities and market imperfections and how to analyze aggregate surplus when compensation must use restricted policy instruments. We end this section with a discussion of how to apply our framework to dynamic and stochastic economies.

5.1 Externalities and Market Imperfections

In this section, we extend Theorem 1 to accommodate market imperfections and externalities. To do so, as in Baqaee and Farhi (2020), we model market imperfections using reduced-form tax-like wedges between price and marginal costs. We then provide a one-good economy example with externalities and study optimal policy.

The household block of the model is unchanged. Each agent chooses consumption and location to maximize utility given prices, wages, and transfers.

We model externalities by allowing for spillovers or congestion in production. Sup-

pose that the production function for i is

$$y_i = z_i F_i(x_{i1}, \ldots x_{iN}, \ell_{i1}^{\alpha_{i1}} l_{i1}, \ldots, \ell_{iR}^{\alpha_{iR}} l_{iR}),$$

where F_i is constant returns to scale, x_{ij} are intermediate inputs, and l_{ir} is efficiency units of labor r used by i. Let $\ell_{ir} = l_{ir}/Z$ denote i's employment of labor r not including the aggregate scaling factor Z. The parameter α_{ir} captures the externality not internalized by producers of i from employing workers of type r. Individual producers take aggregate employment ℓ_{ir} as given when choosing inputs. If $\alpha_{ir} > 0$, then the externality is positive and if $\alpha_{ir} < 0$, it is negative.[21]

Producers minimize the cost of producing their output, taking input prices and the externality as given. Let mc_i denote (private) marginal cost of i and suppose that the price of good i is given by

$$p_i = \mu_i \text{mc}_i, \tag{23}$$

where $\mu_i > 0$ is a tax, markup, or any other wedge between price and marginal cost for producer i.[22] Lump-sum transfers across households must equal total profits and net revenues generated by the wedges,

$$\int T_h dh = \sum_i p_i y_i \left(1 - \frac{1}{\mu_i}\right). \tag{24}$$

In the decentralized equilibrium, transfers are distributed across households according to some rule. We now define an equilibrium with wedges.

Definition 5 (Equilibrium with Wedges)**.** An equilibrium is a collection of consumptions, c_h, location choices, l_h, transfers, $\{T_h\}$, outputs, y_i, input choices, $\{x_{ij}, l_{ir}\}$, prices, p_i, and wages, w_r, such that each agent chooses consumption and location to maximize utility subject to their budget constraint, producers choose inputs to minimize costs given their output and taking prices and the externality terms as given, prices satisfy (23), transfers add up to net revenues generated by the wedges as in (24), and all markets clear (6) and (7).

[21]The fact that the externality applies to ℓ_{ir}, which is employment not including the scaling factor Z, means that scaling Z, holding the allocation constant, does not mechanically trigger the externality. If workers have homogeneous skills, then ℓ_{ir} is the number of workers of type r employed by i, rather than the efficiency units of labor employed.

[22]Even though we assume that wedges are on gross output only, this is without loss of generality. This is because we can recreate buyer-seller specific wedges by treating firm i's purchases of an input from j as a distinct good (made linearly using j's output). A wedge on this good is then isomorphic to a buyer-seller wedge.

The baseline competitive equilibrium obtains when $\boldsymbol{\mu} = \mathbf{1}$ and $\alpha_{ir} = 0$ for every i and r. Define the set of feasible allocations given productivity shifters $\boldsymbol{z}$, wedges $\boldsymbol{\mu}$, and aggregate factor-augmenting productivity Z to be

$$\mathcal{C}(\boldsymbol{z}, \boldsymbol{\mu}, Z) \equiv \{\{c_h, l_h\}_h \text{ supported via equilibrium given } \boldsymbol{z}, \boldsymbol{\mu}, Z, \text{ and some transfers.}\}$$

Aggregate surplus A is defined exactly as in Equation (9) but using $\mathcal{C}(\boldsymbol{z}, \boldsymbol{\mu}, Z)$ in place of the set of feasible allocations in the competitive equilibrium.

We can calculate A using a compensated equilibrium with wedges, extending the results in Section 3.2. In the compensated equilibrium with wedges, every agent receives a transfer that keeps them indifferent, the sum of transfers equals total profits and net revenues generated by the wedges, and TFP is contracted by A. Compensated choices, $l_h^{\text{comp}}(Z\boldsymbol{w}, \boldsymbol{p}^c)$ and the net expenditure function, $T_h(Z\boldsymbol{w}, \boldsymbol{p}^c)$ are obtained using Proposition 1, and these individual choices are aggregated to get location-level variables, $L_r^{\text{comp}}(Z\boldsymbol{w}, \boldsymbol{p}^c)$ and $E_r^{\text{comp}}(Z\boldsymbol{w}, \boldsymbol{p}^c)$, exactly as in the baseline model without wedges. The generalization of Theorem 1 is:

Theorem 3 (Aggregate Welfare with Wedges)**.** *Aggregate surplus in TFP-equivalent terms, A, solves*

$$\sum_r E_r^{comp}(\boldsymbol{w}/A, \boldsymbol{p}^c) = \sum_r (w_r/A) L_r^{comp}(\boldsymbol{w}/A, \boldsymbol{p}^c) + \sum_i p_i y_i \left(1 - \frac{1}{\mu_i}\right), \tag{25}$$

where prices clear the modified market clearing conditions

$$E_r^{comp}(\boldsymbol{w}/A, \boldsymbol{p}^c) = p_r^c y_{c(r)} \text{ if } r \in R \qquad \text{and} \qquad p_i \sum_{j \in C+N} x_{ji} = p_i y_i \text{ if } i \in N. \tag{26}$$

and wages clear the modified labor market clearing conditions

$$w_r \left(\sum_{i \in C+N} l_{ir} \right) = Z w_r L_r^{comp}(\boldsymbol{w}/A, \boldsymbol{p}^c). \tag{27}$$

In these market clearing conditions, the firm side consisting of intermediate input demand, factor demand, and production technologies including externalities is the same as in the decentralized equilibrium. All quantities and costs are evaluated in the compensated equilibrium, where $Z = 1/A$.

The proof follows the same steps as in Theorem 1, using the transfer-budget condition (24) in place of zero aggregate transfers.

We use the one-good economy to show how the framework discussed above can incorporate agglomeration externalities.

Example 6 (One-Good Economy with Externalities). Consider again the one-good economy described in Example 1. There is one freely traded consumption good produced from labor, and production in location r is given by

$$y_r = z_r \ell_r^{\alpha} l_r,$$

where z_r is a parameter and ℓ_r is the measure of agents in location r. Since type-r labor is only employed by producers of r, the measure of agents employed by r is the same as the supply of workers of that type: $\ell_r = L_r$. The baseline model, without externalities, has $\alpha = 0$. Combining labor supply and demand, $l_r = ZL_r$, production in location r is

$$y_r = Z z_r L_r^{\alpha} L_r.$$

The wage per efficiency unit, in the decentralized equilibrium, is $w_r = z_r L_r^{\alpha}$, which is now endogenous because of the externality.

Applying Theorem 3 yields the following expression for A:

$$A = \frac{\sum_r w_r L_r^{\text{comp}}(\boldsymbol{w}/A)}{\sum_r E_r^{\text{comp}}(\boldsymbol{w}/A)} = \frac{\sum_r \int w_r \mathbf{1}[l_h^{\text{comp}} = r] dh}{\sum_r \int \bar{c}_{hr} \mathbf{1}[l_h^{\text{comp}} = r] dh}, \tag{28}$$

where

$$w_r = z_r \left(L_r^{\text{comp}}\right)^{\alpha}.$$

The same solution algorithm as before applies, with an extra fixed point, where the local wage w_r depends on compensated labor supply because of the externality.[23]

Since Theorem 3 shows that aggregate welfare can be represented as an aggregate TFP shock in a compensated equilibrium, we can port the analyses of distorted economies with continuous choice, like the results of Petrin and Levinsohn (2012) or Baqaee and Farhi (2020), into distorted discrete choice economies. For example, we can generalize the Hulten-style result in Corollary 2 beyond the perfectly competitive case.

[23] If there are multiple solutions, we must use the one that implies the highest value of A.

Proposition 5 (Hulten's Theorem with Distortions)**.** *The total derivative of aggregate welfare, A, in response to changes in productivities $d\log \boldsymbol{z}$ and wedges $d\log \boldsymbol{\mu}$ is*

$$(1-\sum_i \lambda_i(1-\frac{1}{\mu_i}))d\log A = \sum_i \frac{\lambda_i}{\mu_i} d\log z_i + \sum_i \lambda_i(1-\frac{1}{\mu_i}) d\log y_i + \sum_{i,r} \lambda_{ir}\alpha_{ir} d\log \ell_{ir},$$

where λ_{ir} is i's wage bill for labor r divided by GDP, $d\log y_i$ is the first-order change in its output, and $d\log \ell_{ir}$ is i's change in employment of labor r. Every variable is evaluated in the compensated equilibrium.

The coefficient on the left is primary factors' share of GDP. The three terms on the right capture direct productivity gains, changes in output interacting with existing wedges, and employment changes that generate externalities. When $\boldsymbol{\mu}^0 = \mathbf{1}$ and $\boldsymbol{\alpha} = \mathbf{0}$, this collapses to Corollary 2. However, in the presence of pre-existing distortions or externalities, this expression shows that how a shock changes aggregate welfare depends on how it interacts with pre-existing distortions and externalities, exactly as in the intuition from classical continuous choice economies discussed by Harberger (1964).

The following example illustrates how Proposition 5 can be used to study counterfactuals, including optimal policy.

Example 7 (One-Good Economy). Consider again the one-good economy described in Example 6. Consider a policymaker who chooses production taxes, μ_r in each location r, and has the objective of maximizing total surplus as measured by A. From Proposition 5, setting the policymaker's first-order condition to zero requires that

$$0 = \sum_r \lambda_r(1-\frac{1}{\mu_r}) d\log y_r + \sum_r \frac{\lambda_r}{\mu_r}\alpha d\log \ell_r,$$

where $d\log y_r$ and $d\log \ell_r$ are changes in quantities of goods and employment in the compensated equilibrium in response to changes in wedges and λ_r/μ_r is the wage bill of r relative to GDP in the compensated equilibrium. The production technology implies that $d\log y_r = (1+\alpha)d\log \ell_r - d\log A = (1+\alpha)d\log \ell_r$ at the optimum, since $d\log A = 0$ at this point. Substituting this in we get the first-order condition:

$$0 = \sum_r \lambda_r \left[(1-\frac{1}{\mu_r})(1+\alpha) + \frac{\alpha}{\mu_r}\right] d\log \ell_r.$$

This condition is satisfied by the Pigouvian taxes or subsidies

$$\mu_r = 1/(1+\alpha).$$

This makes the wage $w_r = (1+\alpha)z_r L_r^{\alpha}$, which equates the wage to the social marginal product of labor and causes firms to internalize the externality. If there is a positive externality, $\alpha > 0$, then production is subsidized, and if the externality is negative, $\alpha < 0$, then it is taxed. If there are no externalities, $\alpha = 0$, then this policymaker does not intervene in the decentralized equilibrium.[24]

5.2 Restricted Compensations

In our baseline, we measure aggregate economic surplus left if winners were to compensate losers using lump-sum transfers. This provides a conceptually clean benchmark, but represents a kind of upper-bound for realizable surplus, since costless lump-sum transfers may not be available. However, our definition of aggregate surplus can easily be extended to cases where individual-level lump-sum transfers are not available. We demonstrate this using place-based compensations.

That is, suppose that instead of individual-specific lump-sum transfers, location-level consumption subsidy rates, denoted by τ_r, can be applied to households that choose location r. In this case, the budget constraint of agent h is now:

$$\sum_r (1-\tau_r) p_r^c c_h \mathbf{1}\left[l_h = r\right] = Z \sum_r w_r a_{hr} \mathbf{1}\left[l_h = r\right] + T,$$

where T is a uniform lump-sum transfer (the same for all households). Budget balance requires that

$$T = -\int \sum_r \tau_r p_r^c c_h \mathbf{1}\left[l_h = r\right] dh. \tag{29}$$

All other conditions defining equilibrium are the same as in Section 2. We can now define the feasible set of allocations that can be supported using these place-based policy instruments:

$\mathcal{C}^{pb}(\boldsymbol{z}, Z) \equiv \{\{c_h, l_h\}_{h\in H}$ supported via equilibrium given $\boldsymbol{z}$, Z, and some consumption subsidies$\}$.

[24] In Section 6, we contrast this simple and intuitive policy with the optimal policies implied by maximizing average utility, which does not typically take this Pigouvian formulation.

Given this feasible set, we can use (9), but replace the set of allocations given lump-sum transfers, $\mathcal{C}(\boldsymbol{z}, Z)$, by the set of allocations given place-based consumption subsidies, $\mathcal{C}^{pb}(\boldsymbol{z}, Z)$.

Definition 6. With place-based compensation, the aggregate surplus generated by the change, in TFP-equivalent terms, is

$$A^{pb} = \max \left\{ Z^{-1} : \{c_h, l_h\}_{h \in H} \in \mathcal{C}^{pb}(\boldsymbol{z}, Z) \text{ and } (c_h, l_h) \succeq_h (c_h^0, l_h^0) \text{ for every } h \right\} \tag{30}$$

where c_h^0 and l_h^0 are the consumption and location of h in the status quo.

To see how this works in practice, consider the one-good economy of Example 1. As we show below, since compensation now varies by location, rather than individual, solving for A^{pb} may actually be much more straightforward. We illustrate this using the example below.

Example 8 (One-Good Economy with Place-Based Compensation). Consider the one-good economy in Example 1 again: there is a single freely-traded consumption good produced linearly from labor, and productivity in each location r is z_r. To calculate A^{pb}, rearrange the budget-balance condition, (29), as

$$Z^{-1} = \frac{\sum_r w_r L_r}{\sum_r p^c c_r L_r} = \frac{\sum_r z_r L_r}{\sum_r c_r L_r}, \tag{31}$$

where c_r denotes per-capita consumption in location r, and we use the fact that the real wage in location r is equal to z_r in this one-good economy. Using equations (30) and (31), we can write

$$A^{pb} = \max_{\boldsymbol{c}} \left\{ \frac{\sum_r z_r L_r(\boldsymbol{c})}{\sum_r c_r L_r(\boldsymbol{c})} : u_h(c_{l_h(\boldsymbol{c})}, l_h(\boldsymbol{c})) \geq u_h^0 \text{ for every } h \in H \right\},$$

where $L_r(\boldsymbol{c})$ is labor supply in location r as a function of the vector of location-level per capita consumption $\boldsymbol{c}$.

If, in the solution to the problem above, for every location $r \in R$, some agent $h \in H$ stays in the same location as in the status quo and all locations are non-empty, then compensating the stayers requires ensuring that consumption per capita in each location is at least as high as in the status quo. Hence, we can replace the H inequalities above

with just R inequalities:

$$A^{pb} = \max_{\boldsymbol{c}} \left\{ \frac{\sum_r z_r L_r(\boldsymbol{c})}{\sum_r c_r L_r(\boldsymbol{c})} : c_r \geq c_r^0 \text{ for every } r \in R \right\}.$$

This is a relatively simple constrained maximization problem given the supply function.

To make it more concrete, suppose that $\boldsymbol{L}(\boldsymbol{c})$ is the isoelastic labor supply function:

$$L_r(\boldsymbol{c}) = \frac{c_r^\theta}{\sum_{r'} c_{r'}^\theta}.$$

Then, it is straightforward to show that the solution to A^{pb} satisfies:[25]

$$c_r^{pb} = \max \left\{ z_r^0, \frac{\theta}{\theta+1} \frac{z_r}{A^{pb}} \right\}$$

If the shocks are sufficiently small, $\frac{z_r/z_r^0}{A^{pb}} \leq \frac{\theta+1}{\theta}$ for every r, then the solution sets consumption in each location equal to its status quo value $c_r^{pb} = z_r^0$. This is achieved by scaling aggregate factor productivity by $1/A^{pb}$ and setting consumption subsidy rates appropriately. Since consumption is equal to its status-quo value in every location, no agent switches locations, and hence every agent is exactly indifferent to the status-quo. However, if region r experiences much faster productivity growth than average — in the sense that $\frac{z_r/z_r^0}{A^{pb}} > \frac{\theta+1}{\theta}$ — then consumption in region r must exceed its status-quo level to incentivize additional workers to move there. In this case, stayers in region r are strictly better off than in the status quo. Given the optimal choices $\boldsymbol{c}^{pb}$ and the implied labor supplies $L_r(\boldsymbol{c}^{pb})$, the place-based aggregate welfare measure A^{pb} is obtained from (31).

To a first-order approximation, A^{pb} and A coincide in a perfectly competitive model if transfers are zero in the initial decentralized equilibrium associated with the status quo.

Proposition 6 (Hulten's Theorem with Place-Based Compensations)**.** *Suppose the status quo has no wedges or externalities and all household transfers are zero. To a first-order approxima-*

[25] Substituting the isoelastic shares into the objective and using $c_r^0 = z_r^0$ gives

$$A^{pb} = \max_{\{c_r \geq z_r^0\}} \frac{\sum_r z_r c_r^\theta}{\sum_r c_r^{\theta+1}}.$$

By Dinkelbach (1967), its optimizer also maximizes $\sum_r \left[z_r c_r^\theta - A^{pb} c_r^{\theta+1} \right]$. This problem is separable across locations. The first-order condition, together with $c_r \geq z_r^0$, gives $c_r^{pb} = \max \left\{ z_r^0, \frac{\theta}{\theta+1} \frac{z_r}{A^{pb}} \right\}$.

tion in $\Delta \log \boldsymbol{z} = \log \boldsymbol{z} - \log \boldsymbol{z}^0$*, the TFP-equivalent welfare change with place-based compensations satisfies*

$$\log A^{pb} \approx \log A \approx \sum_i \lambda_i^0 \Delta \log z_i,$$

where λ_i^0 *is the sales-to-GDP ratio of* i *in the decentralized equilibrium at the status quo.*

Intuitively, place-based compensations introduce distortions into the economy. However, the social costs of these distortions are second-order starting at an efficient equilibrium. Hence, to a first-order, A and A^{pb} coincide.

5.3 Dynamic Stochastic Economies

We end this section with a brief description of the relevance of our framework for dynamic stochastic economies.[26] With preferences defined over complete time- and state-contingent plans and complete contingent markets, the compensated-equilibrium construction and Hulten formulas apply using date-0 Arrow-Debreu prices to value sales and GDP. In other words, we expand the commodity space and allow commodities and locations to be state-contingent streams. So, for example, l_h is now a potentially state-contingent vector of location choices over h's entire life. With this relabeling, the results in Section 3 and Section 4 apply to dynamic stochastic economies without change. If the dynamic stochastic economy has incomplete financial markets, then one can use the results in Section 5.1, where financial market incompleteness is captured via household-time-state-specific wedges as in settings with continuous choice (see, e.g., Baqaee and Burstein, 2025b).

6 Comparison to Other Aggregate Measures

In this section, we compare our measure of aggregate welfare with three alternative aggregate measures that are popular in the literature. In Section 6.1, we relate our measure to real GDP as measured in the national accounts. In Section 6.2, we compare our results to the sum of compensating variations, previously analyzed by Small and Rosen (1981) in a discrete-choice context. Finally, in Section 6.3, we contrast our measure with the popular utilitarian or average utility measure frequently used in the spatial economics literature. We show, using theory and quantitative examples, that these alternative measures have serious flaws.

[26] See, for example, Dix-Carneiro (2014) and Caliendo et al. (2019) for dynamic discrete-choice models of trade and labor reallocation.

6.1 Real GDP

We contrast changes in aggregate welfare, measured via A, with changes in real GDP (or real consumption), as measured in the national accounts.

Definition 7 (Real GDP)**.** To a first-order approximation, the change in real GDP is the share-weighted sum of changes in final consumption quantities:

$$\log Y = \sum_{r \in R} \frac{p^0_{c(r)} C^0_r}{\sum_{r'} p^0_{c(r')} C^0_{r'}} \Delta \log C_r,$$

where $C_r = \int c_h \mathbf{1}[l_h = r] dh$ is total consumption in location r.[27]

Using Hulten (1978), we can write real GDP as:

$$\log Y \approx \sum_{i \in C+N} \lambda^0_i \Delta \log z_i + \sum_r \frac{w^0_r}{\sum_{r'} w^0_{r'} L^0_{r'}} \Delta L_r.$$

Holding factor supplies fixed, real GDP responds positively to productivity growth. Holding productivities fixed, real GDP benefits from relocations of labor from lower-wage towards higher-wage locations. These labor relocations, ΔL_r, are endogenous and require solving the general equilibrium model.

Note that unlike A, real output does not respect the Pareto principle. A shock can make every household better off, and cause real GDP to decline. To see this, consider a positive productivity shock to a low-wage location. In this case, output can fall if the shock induces large flows out of high-wage locations into the still-low-wage location.[28] Intuitively, because agents care about amenities, they may move to a lower-wage region and this reduces production even though every agent is made (weakly) better off.

In contrast, aggregate welfare, A, does not respond to relocations to a first order (see Corollary 2). If households choose to move from high- to low-wage locations, this has no first-order effect on aggregate welfare. The change in wages experienced by movers

[27] Technically, different definitions of real GDP using different index number formulas (e.g. Laspeyres, Paasche, Tornqvist, Divisia, etc.) may disagree with each other at higher orders of approximation. However, they all coincide to a first-order approximation, and for our purposes in this section, a first-order approximation suffices.

[28] For an explicit example, consider the one-good economy with isoelastic labor supply, as in Example 5, the elasticity of real output with respect to a productivity shock is

$$\frac{\Delta \log Y}{\Delta \log z_r} \approx \frac{w^0_r L^0_r}{\sum_k w^0_k L^0_k} + \theta \left[\frac{w^0_r}{\sum_k w^0_k L^0_k} - 1 \right] L^0_r.$$

This elasticity can be negative if θ is high and w_r relatively low.

is exactly offset by the change in the amenity value of the move. Hence, whereas real output changes due to the changes in location choices, aggregate welfare does not. We summarize this in the proposition below.

Proposition 7 (First-Order Difference between Real GDP and Welfare)**.** *To a first-order approximation, the difference between real GDP and TFP-equivalent aggregate surplus is given by the change in income caused by relocation:*

$$\Delta \log Y - \Delta \log A \approx \sum_r \frac{w_r^0}{\sum_{r'} w_{r'}^0 L_{r'}^0} \Delta L_r.$$

Recall that multifactor productivity is defined to be the difference between real GDP growth and the growth in quality-adjusted labor inputs:

$$\Delta \log A^{MFP} \approx \Delta \log Y - \sum_r \frac{w_r^0}{\sum_{r'} w_{r'}^0 L_{r'}^0} \Delta L_r.$$

Hence, Proposition 7 implies that, in perfectly competitive economies, aggregate welfare is correctly measured by multifactor productivity growth to a first-order approximation: $\Delta \log A^{MFP} \approx \Delta \log A$. That is, once we separate changes in real GDP due to technology shocks from changes due to relocation, then multi-factor productivity, as measured by a national income accountant, coincides to a first-order with our measure of aggregate welfare. This equivalence can break down beyond a first-order approximation.

We close with an important and tractable special case. Suppose there is no amenity value associated with different locations, and that there is a single consumption good. Agents can have idiosyncratic skills by location, so that a_{hr} can vary as a function of both h and r. In this case, real GDP, Y, and our measure of aggregate welfare, A, coincide. These assumptions hold in many models of occupational choice, but are less likely to hold in spatial models.

Corollary 5 (Coincidence of real GDP and aggregate welfare)**.** *Suppose there is a common consumption good (consumed in every location), and that $\succeq_h$ ranks (c_h, l_h) only by c_h. Then*

$$A \approx Y.$$

This follows from the fact that under these assumptions, every household chooses location to maximize nominal income, so marginal relocations do not contribute to the change in maximized nominal income (by the envelope theorem), which implies that

$\log Y = \log A^{MFP}$. The result then follows from Proposition 7.[29]

6.2 Sum of Compensating Variations

In partial equilibrium contexts, where there is an outside good, sometimes called "money", a common measure of aggregate welfare is the sum of compensating variations (see, e.g., Small and Rosen, 1981). Using our notation, this measure is:

$$S \equiv -\overline{T}(\boldsymbol{w}, \boldsymbol{p}^c) = -\int T_h(\boldsymbol{w}, \boldsymbol{p}^c) dh,$$

where wages and prices, $\boldsymbol{w}$ and $\boldsymbol{p}^c$, are evaluated in the counterfactual decentralized equilibrium instead of the compensated equilibrium. In words, this is the amount of money left, in terms of the numeraire, after winners from a shock exactly compensate the losers, holding prices and wages constant at their new values in the decentralized equilibrium.[30]

We discuss this measure in more detail in Appendix D.2. If prices and wages are exogenous (e.g. as in partial equilibrium settings), then S and A are very similar, since one is defined by $S + \overline{T}(\boldsymbol{w}, \boldsymbol{p}^c) = 0$ and the other is defined by $\overline{T}(\boldsymbol{w}/A, \boldsymbol{p}^c) = 0$. Indeed, when prices and wages are exogenous, it can be shown that S measures an improvement if, and only if, A measures an improvement.

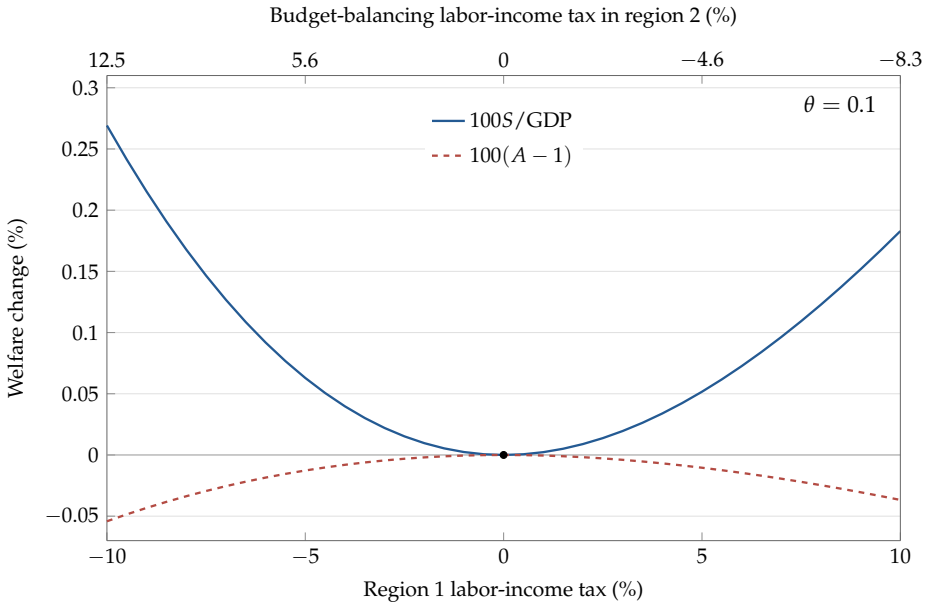
However, when wages and prices are endogenous, the sum of compensating variations, S, can mistake redistribution for efficiency gains. It values gains and losses at the new prices, but compensating losers can itself change relative prices and completely eliminate an apparent surplus. As the following example shows, maximizing this measure can therefore favor extreme redistribution even when it creates no efficiency gains.

Example 9 (Maximizing the sum of compensating variations). Consider an economy with two locations, and households are evenly split living across the two locations. Labor supply is isoelastic, as in Example 5, with elasticity θ. In each location r, a competitive firm produces good r one-for-one from local labor and goods are freely traded. The consumption good in each region is Cobb–Douglas with a home share of $\beta > 1/2$.

[29] If we define real GDP using a chained Divisia index, then Corollary 5 can be strengthened into a global statement. That is, under the assumptions of Corollary 5, a Divisia measure of real GDP nonlinearly coincides with aggregate welfare A.

[30] An unimportant difference between the sum of compensating variations and A is that the former is measured in dollars, whereas A is unit-free, being defined as a ratio. In Appendix D.2 we show that the conclusions of this section do not hinge on this distinction.

Figure 1: Distorting redistributive taxes raise S but lower A.



Notes: The mobility elasticity is $\theta = 0.1$, the local-good share is $\beta = 3/4$. The lower axis varies the labor-income tax in location 1; the upper axis shows the budget-balancing tax in location 2. Negative tax rates are subsidies.

Suppose that a policymaker has access to place-based labor-income tax rates τ_1 and τ_2, where a negative rate is a subsidy. Household h's after-tax income is $(1 - \tau_{l_h})w_{l_h}$, and government budget balance requires $\tau_1 w_1 L_1 + \tau_2 w_2 L_2 = 0$.[31] The competitive equilibrium with $\tau_1 = \tau_2 = 0$ is Pareto efficient, so any change in the taxes is purely redistributive to a first-order, and causes distortions to second order.

Figure 1 shows how the sum of compensating variations, S, and our measure of aggregate welfare, A, change in response to a labor tax on region 1 (with an off-setting tax or subsidy on the other region to maintain budget balance). The policymaker that seeks to maximize the sum of compensating variations, S, prefers to tax one region and subsidize the other because this causes the sum of compensating variations to become positive, even though this policy raises inequality and causes the equilibrium to become inefficient. The apparent surplus arises because redistribution lowers the relative price of the good favored by the loser. At those new prices, compensating the loser looks inexpensive and so the sum of compensating variations is strictly positive. However, actually carrying out compensation would change relative prices and, since the labor taxes are distorting, in fact result in negative surplus. Our measure of aggregate welfare, A, correctly records

[31] In calculating S, we evaluate T_h at the after-tax wages $\{(1 - \tau_r)w_r\}$, holding wages, consumption prices, and tax rates fixed.

a decline in aggregate surplus if the policymaker chooses to use the distorting income taxes.[32]

6.3 Average Utility

We end this section by comparing our measure to aggregate welfare measured using the cross-sectional average of utilities. In the spatial economies literature, this is by far the most common way to evaluate aggregate welfare. We begin by defining this measure and proving a non-identification result: the change in welfare according to an average utility metric is not pinned down by ordinal preference relations. Next, we illustrate that these issues matter quantitatively using a version of the Allen and Arkolakis (2014) economy with heterogeneous tastes.

6.3.1 Definition and Non-Identification

Definition 8. Given real wages across locations, define average utility as

$$W(\boldsymbol{w}/\boldsymbol{p}^c) = \int_h \left[\max_{l_h \in R} u_h\left(c_{h,l_h}, l_h\right)\right] dh = \mathbb{E}\left[\max_{l_h \in R} u_h\left(a_{hl_h} w_{l_h}/p^c_{l_h}, l_h\right)\right], \tag{32}$$

where $\mathbb{E}$ denotes the population-weighted cross-sectional average and $a_{hl_h} w_{l_h}/p^c_{l_h}$ is household h's consumption if location l_h is chosen. Although (32) is sometimes called "expected utility," the expectation is across households, not over lotteries faced by any household, so the label is misleading.[33]

Define the consumption-equivalent variation for a utilitarian social welfare function to be the scalar A^W that solves:

$$W(\boldsymbol{w}/(\boldsymbol{p}^c A^W)) = W(\boldsymbol{w}^0/\boldsymbol{p}^{c,0}). \tag{33}$$

[32] This phenomenon also occurs in general equilibrium models with continuous choice, where it is known that pure redistributions can cause the sum of CVs to become positive (Boadway, 1974). See Baqaee and Burstein (2025a) for more information in a continuous choice setting.

[33] A von Neumann–Morgenstern expected utility function represents an ordinal preference relation over lotteries of allocations; see Mas-Colell et al. (1995, Chapter 6). In this setting, households' preferences are defined over consumption-location bundles given preference parameters. Utility functions do not rank one set of preference parameters, belonging to one household, against another set of preference parameters, belonging to another household. A cross-sectional average of utilities is an expectation of utility values across households with different preference parameters, so it is not formally a von Neumann–Morgenstern expected utility function. For a discussion of these issues in a dynamic model where households have state-dependent preferences, see Appendix D.4.

The object in (32) is a social welfare function. It requires cardinal interpersonal utility comparisons that are not disciplined by observed choices. Unlike our measure of aggregate welfare A, this means that A^W is not invariant to increasing transformations of individual utility functions and that its value is not pinned down by observables. The following proposition formalizes this idea.

Proposition 8 (Arbitrariness of average utility)**.** *Hold fixed the status-quo and counterfactual wage and price vectors, households' ordinal preferences $\{\succeq_h\}$, and hence their individual compensating variations A_h defined in* (10)*. Then A^W can be made equal to any value in the interval*

$$\left(\min_{h\in H} A_h, \max_{h\in H} A_h\right),$$

depending on the cardinal utility representations, and always lies between these endpoints.

This proposition shows that A^W can take any value inside the range spanned by individual compensating variations, depending on the choice of utility function representation.[34] Thus, if the shock creates both winners and losers, then whether average utility rises, falls, or stays the same depends on untestable assumptions about how ordinal preferences are mapped into cardinal utilities.

One might imagine that if we restrict ourselves to a priori "reasonable" cardinalizations that treat households symmetrically, then A^W may be pinned down, but this is not true. The following example uses the popular isoelastic location-choice specification to show that A^W varies erratically across alternative, seemingly reasonable choices of utility cardinalizations. Of course, all cardinalizations are observationally equivalent and consistent with the same underlying preference relations.

6.3.2 Evidence from the Allen-Arkolakis Economy

To illustrate this issue, we use a version of the Allen and Arkolakis (2014) model, augmented with taste heterogeneity. We show that the average utility gain associated with shocks and policies is affected by arbitrary and untestable assumptions mapping ordinal preference rankings into numerical utility values.

Sketch of model. Each location, r, produces a differentiated good using labor with productivity $\bar{z}_r$. Goods from r' sold to r are subject to iceberg cost $\tau_{rr'}$. Consumers in each

[34] Under some mild conditions—satisfied, for example, in the standard case where ϵ_{hr} have extreme value distributions—the lower and upper bounds in Proposition 8 equal the smallest and largest proportional changes in regional real wages between the status quo and the counterfactual equilibrium.

location have CES preferences over goods from different origins with elasticity of substitution σ. The price of the consumption bundle in location r and the expenditure share in location r on goods produced in region r' are

$$p_r^c = \left[\sum_{r'} \left(\frac{\tau_{rr'} w_{r'}}{\bar{z}_{r'}}\right)^{1-\sigma}\right]^{1/(1-\sigma)}, \qquad \lambda_{rr'} = \frac{\left(\tau_{rr'} w_{r'} / \bar{z}_{r'}\right)^{1-\sigma}}{\sum_{r''} \left(\tau_{rr''} w_{r''} / \bar{z}_{r''}\right)^{1-\sigma}}.$$

Trade balance between locations implies that

$$w_{r'} L_{r'} = \sum_r \lambda_{rr'} w_r L_r.$$

Productivity may be subject to an externality that depends on regional population and is governed by a parameter α:

$$\bar{z}_r = z_r L_r^{\alpha}.$$

The decentralized equilibrium is Pareto efficient if $\alpha = 0$.

Heterogeneity in tastes is modeled using the preferences in Example 5. Agent h has utility function

$$u_h(c_r, r) = f_h\big(B_r \varepsilon_{hr} c_r\big),$$

where ε_{hr} is distributed according to a Fréchet distribution with shape parameter θ, B_r is a parameter that controls the desirability of location r, and f_h is any arbitrary increasing function. Typically, the literature assumes that $f_h(x) = x$ so that

$$u_h(c_r, r) = B_r c_r \varepsilon_{hr}.$$

Under this popular cardinalization, A^W has a familiar closed-form solution:

$$A^W = \left[\frac{\sum_r (B_r c_r)^{\theta}}{\sum_r (B_r c_r^0)^{\theta}}\right]^{1/\theta}, \tag{34}$$

where $c_r = w_r / p_r^c$ and $c_r^0 = w_r^0 / p_r^{c,0}$ are consumption in the counterfactual and status-quo equilibria. This is a widely used measure of aggregate welfare in spatial models.

However, other cardinalizations are equally valid. For example, suppose we write

$$u_h(c_r, r) = B_r c_r \frac{\varepsilon_{hr}}{\max_{r'} \varepsilon_{hr'}}. \tag{35}$$

This has no implication for choices, but normalizes each agent's vector of location-specific tastes so that its highest value is one and all other values are below one. One can argue

that this is a more "reasonable" cardinalization of utility because it does not allow for the possibility that some agents have high tastes in every location while others have low tastes in every location. Under this cardinalization, however, A^W does not obey (34).

Regardless of the choice of utility cardinalization, labor supplies take the form:

$$L_r = \frac{(B_r w_r / p_r^c)^\theta}{\sum_{r'} (B_{r'} w_{r'} / p_{r'}^c)^\theta}. \tag{36}$$

Calibration. We follow the calibration strategy in Allen and Arkolakis (2014). The model has 3,109 counties in the continental United States ($R = 3,109$), and the Armington elasticity of substitution is $\sigma = 9$. We start by considering the no externalities case, $\alpha = 0$. For illustration, we use a baseline mobility elasticity value of $\theta = 3$. We use the trade-cost matrix $\tau_{rr'}$ in Allen and Arkolakis (2014), which is estimated using geographic data on roads, railways, and navigable waterways, together with Commodity Flow Survey trade flows.[35] We calibrate z_r and B_r so that the model exactly reproduces the same initial population and expenditure shares by county. For any given shock or policy, different cardinalizations of utility use the same primitives and calibration and generate identical prices, wages, migration, and trade flows. Only the reported average-utility welfare measure changes.

Response of aggregate welfare to regional productivity changes. To show the different ways that A^W responds to productivity shocks depending on cardinal assumptions, we shock the productivity of each county r and report the response of A^W compared to the Hulten benchmark:

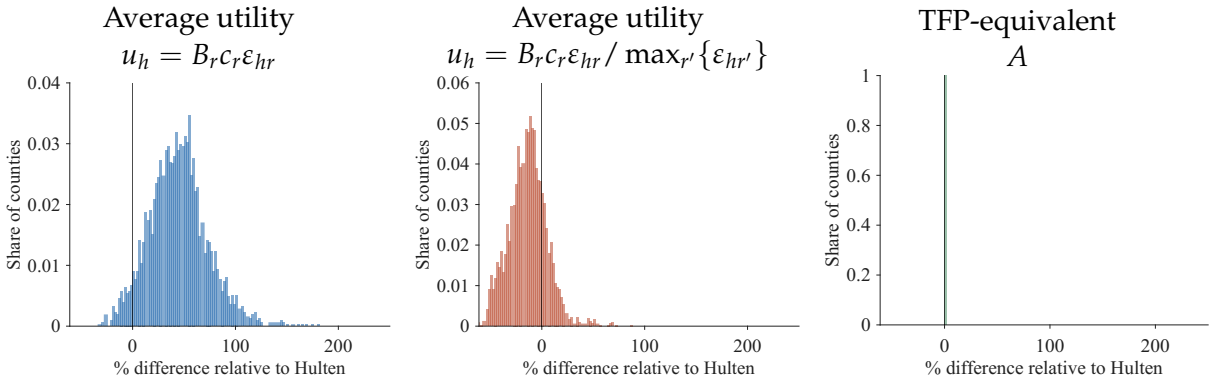
$$100 \times \frac{\partial \log A^W / \partial \log z_r - \lambda_r}{\lambda_r}.$$

Figure 2 plots this statistic across counties for the two different cardinalizations of utility. This statistic is also reported by Donald et al. (2026), using the standard non-normalized cardinalization of utility $c_r \varepsilon_{hr}$. As they point out, the response of average utility does not obey Hulten's theorem. In their language, this is because of "the spatial dispersion in marginal utility." However, since the level of marginal utility across individuals is not pinned down by observables (only the ratio of marginal utilities across goods for each individual is pinned down by observables), the deviations from Hulten are driven by the choice of utility cardinalization. The second panel illustrates this by changing the cardi-

[35] Allen and Arkolakis (2014) study the aggregate welfare loss of losing access to the Interstate-Highway-System, which is modeled as an increase in trade costs. In Appendix D.3, we show that, once we allow for heterogeneous preferences, the change in aggregate welfare using an average utility associated with a counterfactual like this depends on the cardinalization of individual utility functions.

nalization and normalizing by the max ε_{hr} for each h. The result is that the deviations from Hulten's theorem differ substantially across the two panels. The final panel on the right shows that our measure of aggregate surplus, A, obeys Hulten's theorem (Corollary 2), as illustrated by the mass concentrated near zero in the histogram.

Figure 2: County productivity shocks and deviations from Hulten's theorem

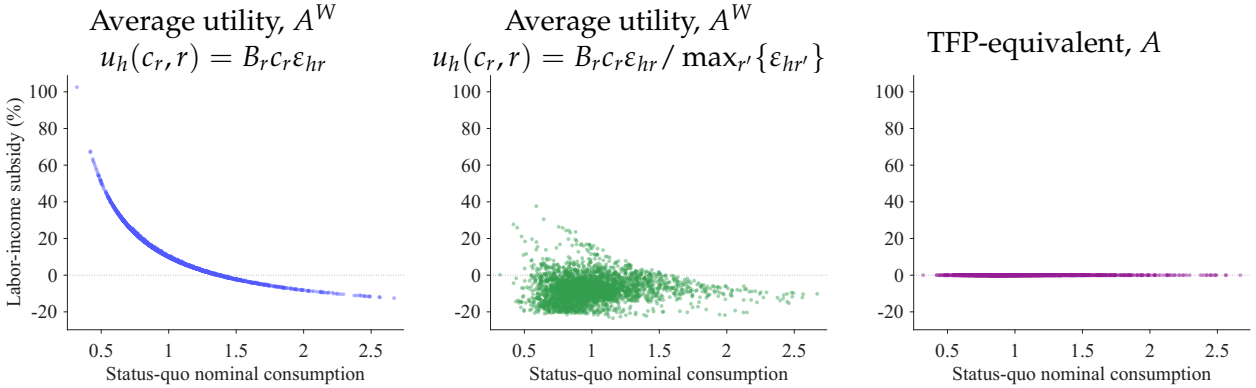


Notes: The first two panels plot the cross-county distribution of $100(\partial \log A^W / \partial \log z_r - \lambda_r)/\lambda_r$ in the competitive specification $\alpha = 0$. The first two panels use observationally equivalent cardinal representations of the same preferences. The third uses TFP-equivalent aggregate surplus. In practice, we approximate these derivatives using log-productivity increases of 0.01 one county at a time.

Optimal place-based policies. Cardinalization also affects policy recommendations. In our no-externality benchmark the competitive equilibrium is Pareto efficient. A policymaker maximizing A therefore leaves every place-based subsidy at zero. A policymaker maximizing average utility instead intervenes in order to pursue redistributive objectives embedded in cardinalizing choices. The first two panels of Figure 3 show that two observationally equivalent cardinalizations imply large and sharply different policies. The third panel maximizes aggregate welfare as measured by A, which selects zero subsidies because the equilibrium is efficient.[36]

[36]For a discussion of the role of social welfare functions in determining the optimality of place-based policies in spatial models, see Fajgelbaum and Gaubert (2025). Relatedly, Mongey and Waugh (2024) point out that redistributive concerns can strongly affect policy prescriptions when average utility is used as the objective. They provide an alternative interpretation where average utility is interpreted as the expected utility of individuals who are ex-ante identical (before they know their tastes) and lack access to insurance markets for their idiosyncratic tastes. In this case, the redistributive desires of the policymaker are interpreted as providing insurance. However, if such an ex-ante state does not exist, then the results of this type of analysis rest on an untestable cardinalizing assumption. That is, if households cannot make choices in the ex-ante stage that reveal their risk preferences about how they rank one set of taste parameters against another, then choice data identifies utility up to a monotone transformation for each h, and the ex-ante expected utility function cannot be recovered from the data. This means that the amount of insurance which is optimal depends on untestable assumptions about how individual utility functions are cardinalized. See also the discussion by Davis and Gregory (2021).

Figure 3: Optimal place-based policies without externalities



Notes: The first two panels maximize average utility under, respectively, $u_h(c_r, r) = B_r c_r \varepsilon_{hr}$ and $u_h(c_r, r) = B_r c_r \varepsilon_{hr} / \max_{r'}\{\varepsilon_{hr'}\}$. The third panel maximizes A. The model has $\alpha = 0$ and $\theta = 3$. Positive values denote subsidies and negative values denote taxes.

Now consider a version of the model with externalities. In this case, the policymaker that maximizes aggregate surplus A does intervene if the initial equilibrium is Pareto inefficient. To see this, suppose instead that we consider the case with positive agglomeration externalities $\alpha > 0$ (recalibrating z_r and B_r to keep population and GDP shares constant). Productivity in region r is

$$\bar{z}_r = z_r L_r^{\alpha}.$$

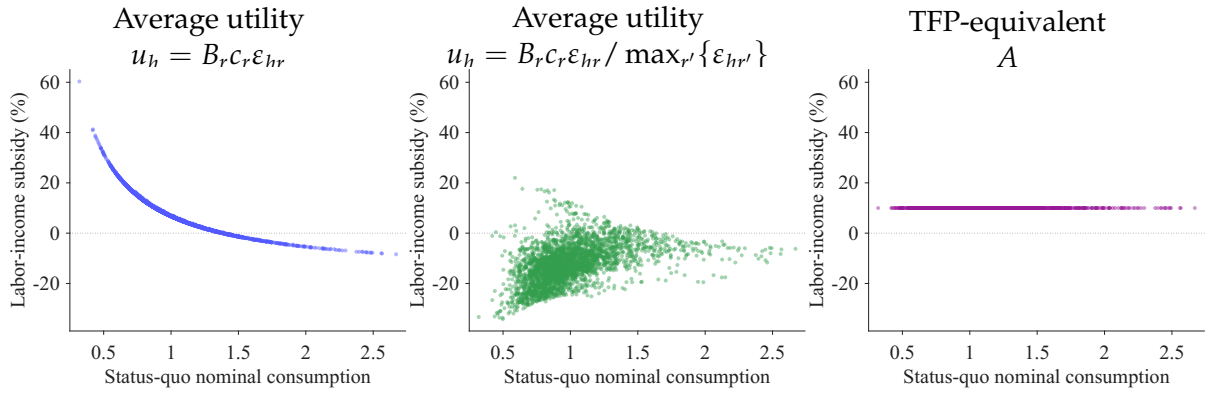
Private producers pay $p_r \bar{z}_r$ per efficiency unit, while the social marginal revenue product is $(1+\alpha) p_r \bar{z}_r$. With a labor-income subsidy τ_r, the worker receives $(1+\tau_r) p_r \bar{z}_r$. Maximizing A therefore delivers the Pigouvian rule

$$\tau_r = \alpha,$$

exactly as in Example 7.

Compare this with the first two schedules in Figure 4, which are county-specific optimal tax rates for different objectives. The policymaker that maximizes A^W responds strongly to cardinal choices about utility and mixes the Pigouvian reasons for intervention with the desire to redistribute. In contrast, the policymaker that maximizes A implements Pigouvian subsidies only.

Figure 4: Optimal place-based policies with an agglomeration externality



Notes: The first two panels maximize average utility under, respectively, $u_h(c_r, r) = B_r c_r \varepsilon_{hr}$ and $u_h(c_r, r) = B_r c_r \varepsilon_{hr} / \max_{r'}\{\varepsilon_{hr'}\}$. The third panel maximizes A. The model has $\alpha = 0.1$ and $\theta = 3$. Positive values denote subsidies and negative values denote taxes.

7 Conclusion

We develop a revealed-preference measure of aggregate economic surplus for general equilibrium economies with discrete choice. We measure the aggregate surplus created by a change in primitives by the amount of factor endowments that can be removed while still ensuring every household can be kept indifferent to the status quo. We show that an aggregate expenditure function summarizes the household side of this calculation. In the undistorted benchmark, the first-order result is especially simple: productivity changes are valued using producers' initial sales relative to GDP. Beyond first order, we provide an exact characterization using compensated supply and demand, together with tractable second-order approximations using observable shares and elasticities.

We contrast our welfare measure with chain-weighted real GDP, the sum of compensating variations, and the cross-sectional average of utilities. We show that these popular alternatives have serious flaws. Real GDP can fall when every household is better off. The sum of compensating variations can rise in response to inefficient redistributive taxes that move the economy inside the Pareto frontier and raise inequality. When a change creates both winners and losers, even the sign of its average-utility welfare effect depends on arbitrary assumptions about how to map each individual's ordinal preferences into interpersonally comparable utility numbers.

While our framework can quantify aggregate surplus generated by changes in technologies and policies, it deliberately leaves the distribution of that surplus as a separate policy question. Our measure can also be applied to specific subpopulations or evaluated under restrictions on compensation. Applying our framework to quantitative dynamic discrete-choice models with incomplete financial markets is a natural next step.

References

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.

Anderson, S. P., A. De Palma, and J. F. Thisse (1992). *Discrete choice theory of product differentiation*.

Bagga, S., L. F. Mann, A. Şahin, and G. L. Violante (2025). Job amenity shocks and labor reallocation. Technical report, National Bureau of Economic Research.

Baqaee, D. and A. Burstein (2025a). Aggregate productivity with heterogeneous agents. Technical report.

Baqaee, D. R. and A. Burstein (2023). Welfare and output with income effects and taste shocks. *The Quarterly Journal of Economics 138*(2), 769–834.

Baqaee, D. R. and A. Burstein (2025b, September). Misallocation due to incomplete markets. Working Paper 34233, National Bureau of Economic Research.

Baqaee, D. R. and E. Farhi (2019). The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem. *Econometrica 87*(4), 1155–1203.

Baqaee, D. R. and E. Farhi (2020). Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics 135*(1), 105–163.

Baqaee, D. R. and E. Farhi (2024). Networks, barriers, and trade. *Econometrica 92*(2), 505–541.

Bhattacharya, D. (2015). Nonparametric welfare analysis for discrete choice. *Econometrica 83*(2), 617–649.

Bhattacharya, D. (2018). Empirical welfare analysis for discrete choice: Some general results. *Quantitative Economics 9*(2), 571–615.

Boadway, R. W. (1974). The welfare foundations of cost-benefit analysis. *The Economic Journal 84*(336), 926–939.

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., M. Dvorkin, and F. Parro (2019). Trade and labor market dynamics: General equilibrium analysis of the china trade shock. *Econometrica 87*(3), 741–835.

Dagsvik, J. K. and A. Karlström (2005). Compensating variation and hicksian choice probabilities in random utility models that are nonlinear in income. *The Review of Economic Studies 72*(1), 57–76.

Daly, A. J. and S. Zachary (1978). Improved multiple choice models. In D. A. Hensher and M. Q. Dalvi (Eds.), *Determinants of Travel Choice*, pp. 335–357. Farnborough, UK: Saxon House.

Davis, M. and J. M. Gregory (2021). Place-based redistribution in location choice models. Technical report, National Bureau of Economic Research.

Debreu, G. (1951). The coefficient of resource utilization. *Econometrica: Journal of the Econometric Society*, 273–292.

Dingel, J. I. and F. Tintelnot (2020). Spatial economics for granular settings. Technical report, National Bureau of Economic Research.

Dinkelbach, W. (1967). On nonlinear fractional programming. *Management science 13*(7), 492–498.

Dix-Carneiro, R. (2014). Trade liberalization and labor market dynamics. *Econometrica 82*(3), 825–885.

Domar, E. D. (1961). On the measurement of technological change. *The Economic Journal 71*(284), 709–729.

Donald, E., M. Fukui, and Y. Miyauchi (2026). Unpacking aggregate welfare in a spatial economy. *Review of Economic Studies*.

Donaldson, D. (2025). Transport infrastructure and policy evaluation. In *Handbook of Regional and Urban Economics*, Volume 6, pp. 287–352. Elsevier.

Dubé, J.-P., J. Joo, and K. Kim (2025). Discrete/continuous choice models and representative consumer theory. *Journal of Economic Theory*, 106012.

Fajgelbaum, P. D. and C. Gaubert (2025). Optimal spatial policies. Technical report, National Bureau of Economic Research.

Fajgelbaum, P. D., E. Morales, J. C. Suárez Serrato, and O. Zidar (2019). State taxes and spatial misallocation. *The Review of Economic Studies 86*(1), 333–376.

Harberger, A. C. (1964). The measurement of waste. *The American Economic Review 54*(3), 58–76.

Harberger, A. C. (1971). Three basic postulates for applied welfare economics: an interpretive essay. *Journal of Economic literature 9*(3), 785–797.

Heise, S. and T. Porzio (2022). Labor misallocation across firms and regions. Technical report, National Bureau of Economic Research.

Hortaçsu, A. and J. Joo (2023). *Structural econometric modeling in industrial organization and quantitative marketing: theory and applications*. Princeton University Press.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2019). The allocation of talent and us economic growth. *Econometrica 87*(5), 1439–1474.

Hulten, C. R. (1978). Growth accounting with intermediate inputs. *The Review of Economic Studies*, 511–518.

Kim, R. and J. Vogel (2020). Trade and welfare (across local labor markets). Technical report, National Bureau of Economic Research.

Kleiner, M. M. and E. J. Soltas (2023). A welfare analysis of occupational licensing in us states. *Review of Economic Studies 90*(5), 2481–2516.

Lamadon, T., M. Mogstad, and B. Setzler (2022). Imperfect competition, compensating differentials, and rent sharing in the us labor market. *American Economic Review 112*(1), 169–212.

Mas-Colell, A., M. D. Whinston, and J. Green (1995). *Microeconomic theory*. Oxford university press New York.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. New York: Academic Press.

McFadden, D. (1981). Econometric models of probabilistic choice. In C. F. Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data with Econometric Applications*, pp. 198–272. Cambridge, MA: MIT Press.

Mongey, S. and M. E. Waugh (2024). Discrete choice, complete markets, and equilibrium. Technical report, National Bureau of Economic Research.

Monte, F., C. Porcher, and E. Rossi-Hansberg (2023). Remote work and city structure. Technical report, National Bureau of Economic Research.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018). Commuting, migration, and local employment elasticities. *American Economic Review 108*(12), 3855–3890.

Petrin, A. and J. Levinsohn (2012). Measuring aggregate productivity growth using plant-level data. *The RAND Journal of Economics 43*(4), 705–725.

Redding, S. J. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics 101*, 148–167.

Redding, S. J. and E. Rossi-Hansberg (2017). Quantitative spatial economics. *Annual Review of Economics 9*(1), 21–58.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of political Economy 90*(6), 1257–1278.

Schulz, K., A. Tsyvinski, and N. Werquin (2023). Generalized compensation principle. *Theoretical Economics 18*(4), 1665–1710.

Small, K. A. and H. S. Rosen (1981). Applied welfare economics with discrete choice models. *Econometrica: Journal of the Econometric Society*, 105–130.

Tsivanidis, N. (2026, February). Evaluating the impact of urban transit infrastructure: Evidence from bogota transmilenio. *American Economic Review 116*(2), 418–63.

Williams, H. C. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and planning A 9*(3), 285–344.

Appendix

A Appendix for Section 3 **A2**
A.1 Proofs for Section 3 . . . . . A2
A.2 Homogeneous preferences and skills A6
A.3 Relation to the social surplus function A9

B Appendix for Section 4 **A15**
B.1 Input-Output Notation for this Appendix A15
B.2 Proofs for Section 4 . . . . . A16
B.3 Second-Order Approximation A17
B.4 Calculating A using a differential-equation system A26

C Appendix for Section 5 **A29**
C.1 Proofs for Section 5 . . . . . A29

D Appendix for Section 6 **A32**
D.1 Proofs for Section 6 . . . . . A32
D.2 Compensating variations in ratios and differences A35
D.3 Welfare effects of removing the Interstate Highway System A38
D.4 Average utility in dynamic discrete choice A39

Throughout the appendix, the status quo is held fixed in the expenditure functions T_h and $\overline{T}$. Where a path is used, $t = 0$ denotes the status quo and $t = 1$ the counterfactual, so that $\boldsymbol{z}(0) = \boldsymbol{z}^0$ and $\boldsymbol{z}(1) = \boldsymbol{z}$. Superscript comp denotes the compensated equilibrium, in which $Z = 1/A(t)$.

A Appendix for Section 3

A.1 Proofs for Section 3

Proof of Proposition 1. The expenditure function is defined to be

$$T_h(\boldsymbol{w}, \boldsymbol{p}^c) = \min_{T_h, c_h, l_h} \{T_h : u_h(c_h, l_h) \geq u_h^0, \text{ and } \sum_r p_r^c c_h \mathbf{1}[l_h = r] \leq \sum_r a_{hr} w_r \mathbf{1}[l_h = r] + T_h\}.$$

Since utility is increasing in consumption, the utility constraint binds at the optimum. Hence, conditional on choosing l_h, consumption must be $\bar{c}_{hl_h}$, where

$$u_h(\bar{c}_{hl_h}, l_h) = u_h^0.$$

Substituting this into the definition gives

$$T_h(\boldsymbol{w}, \boldsymbol{p}^c) = \min_{T_h, l_h} \{T_h : \sum_r p_r^c \bar{c}_{hr} \mathbf{1}[l_h = r] \leq \sum_r a_{hr} w_r \mathbf{1}[l_h = r] + T_h\}.$$

The budget constraint also binds at an optimum (otherwise the transfer could be reduced), so substituting out for T_h yields

$$T_h(\boldsymbol{w}, \boldsymbol{p}^c) = -\max_{l_h} \{\sum_r a_{hr} w_r \mathbf{1}[l_h = r] - \sum_r p_r^c \bar{c}_{hr} \mathbf{1}[l_h = r]\}.$$

The maximizer of this problem (or a set if there are multiple solutions) is the compensated choice $l_h^{\text{comp}}(\boldsymbol{w}, \boldsymbol{p}^c)$, which satisfies the first expression in Proposition 1. Evaluating the transfer at that choice gives the second expression in the proposition. □

Proof of Proposition 2. By Proposition 1, household h's expenditure function can be written as

$$T_h(\boldsymbol{w}, \boldsymbol{p}^c) = \min_{r \in R} \left\{ p_r^c \bar{c}_{hr} - a_{hr} w_r \right\}.$$

Except for the measure-zero set of households that are indifferent between locations, the envelope theorem therefore implies

$$\frac{\partial T_h}{\partial w_r} = -a_{hr} \mathbf{1}[l_h^{\text{comp}} = r], \qquad \frac{\partial T_h}{\partial p_r^c} = \bar{c}_{hr} \mathbf{1}[l_h^{\text{comp}} = r].$$

Integrating the first equality across households gives

$$\frac{\partial \overline{T}}{\partial w_r} = -\int_h a_{hr} \mathbf{1}[l_h^{\text{comp}} = r] dh = -L_r^{\text{comp}}.$$

Similarly,

$$\begin{aligned}\frac{\partial \overline{T}}{\partial \log p_r^c} &= p_r^c \int_h \bar{c}_{hr} \mathbf{1}[l_h^{\text{comp}} = r] dh \\ &= w_r L_r^{\text{comp}} + \int_h T_h \mathbf{1}[l_h^{\text{comp}} = r] dh = E_r^{\text{comp}},\end{aligned}$$

where the second equality uses $p_r^c \bar{c}_{hr} = a_{hr} w_r + T_h$ for a household whose compensated choice is r. Finally, the expression for T_h above is homogeneous of degree one in $(\boldsymbol{w}, \boldsymbol{p}^c)$, so the aggregate expenditure function is also homogeneous of degree one. Euler's theorem and the two derivatives above give

$$\begin{aligned}\overline{T} &= \sum_r w_r \frac{\partial \overline{T}}{\partial w_r} + \sum_r p_r^c \frac{\partial \overline{T}}{\partial p_r^c} \\ &= \sum_r \left(E_r^{\text{comp}} - w_r L_r^{\text{comp}} \right),\end{aligned}$$

as desired. □

Proof of Theorem 1. Recall that $\mathcal{C}(\boldsymbol{z}, Z)$ is the set of allocations supported by a competitive equilibrium given $\boldsymbol{z}$, Z, and some transfers, and that

$$A = \max \left\{ Z^{-1} : \{c_h, l_h\}_{h \in H} \in \mathcal{C}(\boldsymbol{z}, Z) \text{ and } (c_h, l_h) \succeq_h (c_h^0, l_h^0) \text{ for every } h \right\}.$$

Let $Z^* = 1/A$. Let $\{c_h^*, l_h^*\}_{h \in H}$ be an allocation attaining this maximum. By the definition of $\mathcal{C}(\boldsymbol{z}, Z^*)$, there is a budget-balanced transfer vector $\boldsymbol{T}^*$ that supports this allocation. Given the maintained assumption that every household can be made exactly indifferent at the solution, $u_h(c_h^*, l_h^*) = u_h^0$. Therefore,

$$c_h^* = u_h^{-1}\left(u_h^0; l_h^*\right) = \bar{c}_{h l_h^*},$$

where the inverse is with respect to consumption, holding location fixed.[37] Let $\boldsymbol{w}$ and $\boldsymbol{p}^c$ denote wages and consumption prices in this equilibrium. The household budget constraint becomes

$$p_{l_h^*}^c \bar{c}_{h l_h^*} = \frac{a_{h l_h^*} w_{l_h^*}}{A} + T_h^*.$$

Hence,

$$T_h^* = p_{l_h^*}^c \bar{c}_{h l_h^*} - \frac{a_{h l_h^*} w_{l_h^*}}{A}.$$

[37] If we impose ordinal additive separability, which is not required for this proof, then $g(c_h^*) + \epsilon_{h l_h^*} = f_h^{-1}(u_h^0)$, so $c_h^* = g^{-1}\left(f_h^{-1}(u_h^0) - \epsilon_{h l_h^*}\right) \equiv \bar{c}_{h l_h^*}$.

Because households are exactly indifferent to the status quo, their chosen locations minimize the transfers required for indifference. Thus, Proposition 1 implies

$$l_h^* = l_h^{\text{comp}}(\boldsymbol{w}/A, \boldsymbol{p}^c), \qquad T_h^* = T_h(\boldsymbol{w}/A, \boldsymbol{p}^c).$$

Summing the household budget constraints over households choosing location r gives

$$\begin{aligned}\int_h p_r^c \bar{c}_{hr} \mathbf{1}[l_h^{\text{comp}} = r] dh &= \frac{w_r}{A} L_r^{\text{comp}}(\boldsymbol{w}/A, \boldsymbol{p}^c) \\ &\quad + \int_h T_h(\boldsymbol{w}/A, \boldsymbol{p}^c) \mathbf{1}[l_h^{\text{comp}} = r] dh \\ &= E_r^{\text{comp}}(\boldsymbol{w}/A, \boldsymbol{p}^c).\end{aligned}$$

Since transfers are budget-balanced,

$$0 = \int_h T_h^* dh = \sum_r E_r^{\text{comp}}(\boldsymbol{w}/A, \boldsymbol{p}^c) - \sum_r \frac{w_r}{A} L_r^{\text{comp}}(\boldsymbol{w}/A, \boldsymbol{p}^c).$$

This proves that the solution to (9) satisfies (11). The remaining conditions in Theorem 1 follow because $\boldsymbol{w}$ and $\boldsymbol{p}$ are equilibrium wages and prices at the optimizing allocation.

Conversely, suppose that A, wages, prices, and quantities satisfy the compensated-equilibrium conditions in Theorem 1. Give each household a transfer

$$T_h = T_h(\boldsymbol{w}/A, \boldsymbol{p}^c), \qquad l_h = l_h^{\text{comp}}(\boldsymbol{w}/A, \boldsymbol{p}^c), \qquad c_h = \bar{c}_{hl_h}.$$

By Proposition 1, households optimize and are indifferent to the status quo. Equation (11) implies that transfers are budget-balanced, while the remaining compensated-equilibrium conditions ensure that firms optimize and markets clear. Hence, $\{c_h, l_h\}_{h \in H} \in \mathcal{C}(\boldsymbol{z}, 1/A)$, so A is feasible in (9). Together, these arguments imply that the largest value of A supported by the compensated-equilibrium conditions equals aggregate welfare as defined in (9). □

Proof of Proposition 4. Under Assumption 1, Proposition 1 gives

$$T_h(\boldsymbol{w}, \boldsymbol{p}^c) = \min_{r \in R} \{p_r^c \bar{c}_{hr} - w_r\}.$$

We first show that the compensating transfers $T_h(\boldsymbol{w}, \boldsymbol{p}^c)$ are bounded by $T^{\min}(\boldsymbol{w}, \boldsymbol{p}^c)$ and

$T^{\max}(\boldsymbol{w}, \boldsymbol{p}^c)$. In the status quo,

$$u_h^0 = \max_r u_h \left(c_r^0, r \right).$$

Since $u_h(c_r^0, r) \leq u_h^0$ and $u_h(c, r)$ is strictly increasing in consumption, $\bar{c}_{hr} \geq c_r^0$ for every r, with equality at household h's status-quo location. Therefore,

$$T_h(\boldsymbol{w}, \boldsymbol{p}^c) \geq \min_r \{ p_r^c c_r^0 - w_r \} = T^{\min}(\boldsymbol{w}, \boldsymbol{p}^c).$$

Evaluating the minimum at household h's status-quo location gives

$$T_h(\boldsymbol{w}, \boldsymbol{p}^c) \leq p_{l_h^0}^c \bar{c}_{h l_h^0} - w_{l_h^0} = p_{l_h^0}^c c_{l_h^0}^0 - w_{l_h^0} \leq T^{\max}(\boldsymbol{w}, \boldsymbol{p}^c).$$

The tail-sum formula for expectations implies

$$\overline{T}(\boldsymbol{w}, \boldsymbol{p}^c) = T^{\min}(\boldsymbol{w}, \boldsymbol{p}^c) + \int_{T^{\min}(\boldsymbol{w}, \boldsymbol{p}^c)}^{T^{\max}(\boldsymbol{w}, \boldsymbol{p}^c)} \Pr(T_h(\boldsymbol{w}, \boldsymbol{p}^c) > x)\, dx. \tag{37}$$

We now show how to obtain the probability inside the integral. By the definition of the expenditure function,

$$T_h(\boldsymbol{w}, \boldsymbol{p}^c) > x \quad \Longleftrightarrow \quad \max_r u_h \left(\frac{w_r + x}{p_r^c}, r \right) < \max_r u_h \left(c_r^0, r \right).$$

Since

$$\widetilde{c}_r(x) = \max \left\{ \frac{w_r + x}{p_r^c}, c_r^0 \right\}$$

and $u_h(c, r)$ is strictly increasing in c,

$$\max_r u_h(\widetilde{c}_r(x), r) = \max \left\{ \max_r u_h \left(\frac{w_r + x}{p_r^c}, r \right), \max_r u_h \left(c_r^0, r \right) \right\}.$$

Ignoring ties, which do not affect the integral, $T_h(\boldsymbol{w}, \boldsymbol{p}^c) > x$ if and only if the household's choice under $\widetilde{\boldsymbol{c}}(x)$ is a location r satisfying

$$c_r^0 > \frac{w_r + x}{p_r^c}.$$

Aggregating across households therefore gives

$$\Pr(T_h(\boldsymbol{w}, \boldsymbol{p}^c) > x) = \sum_r L_r(\widetilde{\boldsymbol{c}}(x))\, \mathbf{1} \left[c_r^0 > \frac{w_r + x}{p_r^c} \right].$$

Substituting this expression into (37) implies (19). □

A.2 Homogeneous preferences and skills

In this appendix, we consider an instructive special case where agents are homogeneous in preferences and skill levels ($\succeq_h$ does not vary by h and $a_{hr} = 1$ for every h and r). Locations may nevertheless differ in their common amenity values. For concreteness, in this section we assume that preferences are ordinally additively separable, so that we can represent every $\succeq_h$ using:

$$u_h(c_h, l_h) = f_h\left(g(c_h) + \epsilon_{l_h}\right),$$

where ϵ_r parameterizes the common amenity value of location r. In this section we assume that decentralized equilibria have no transfers. Characterizing aggregate surplus A in this special case is as easy as solving the decentralized equilibrium.

Proposition 9 (Aggregate Welfare with Homogeneous Agents)**.** *Suppose $\succeq_h$ does not vary by h and $a_{hr} = 1$ for every h and r. Define $c_r(\boldsymbol{z}, Z)$ and $L_r(\boldsymbol{z}, Z)$ to be consumption per capita and labor supply in location r in a decentralized equilibrium given productivity vector $\boldsymbol{z}$ and aggregate factor-augmenting productivity Z. Suppose that at least one location is nonempty both in the status quo and in the compensated equilibrium. Then aggregate welfare A satisfies*

$$c_r(\boldsymbol{z}^0, 1) = c_r(\boldsymbol{z}, \frac{1}{A}), \tag{38}$$

for every such location r.

If, in addition, $g(c_h) = \log(c_h)$, then A equals the proportional change in per capita consumption (equivalently, the real wage) between the counterfactual and status quo decentralized equilibria:

$$A = \frac{c_r(\boldsymbol{z}, 1)}{c_r(\boldsymbol{z}^0, 1)}, \tag{39}$$

for any region r that is nonempty in both equilibria.

Before presenting the proof, we provide some intuition. Since agents are homogeneous, in a decentralized equilibrium agents are indifferent among occupied locations. If we scale aggregate factor-augmenting productivity by $1/A$, every agent is affected in the same way — there are no relative winners or losers — and lump-sum transfers are not needed for compensations. If a location r is non-empty in the status quo and in the allocation that solves A, indifference to the status quo requires that A satisfies

$$f_h\big(g(c_r(\boldsymbol{z}, 1/A)) + \epsilon_r\big) = f_h\big(g(c_r(\boldsymbol{z}^0, 1)) + \epsilon_r\big).$$

Because f_h is strictly increasing, it can be removed from both sides. The common amenity ϵ_r then cancels, and strict monotonicity of g implies (38). In words, A must be such that real consumption per capita in each such location r is equated to its status quo value. This shows that computing A is as simple as computing the decentralized equilibrium with different productivity parameters (i.e. we do not have to worry about compensating transfers when agents are homogeneous).

In addition, if $g(c) = \log(c)$, then scaling consumption in every location by the same factor leaves the location choice problem of every agent unchanged. Since agents' location choice problems do not respond to a uniform scaling of consumption, they also do not respond to a uniform scaling of factor-augmenting productivity. Hence, $c_r(\boldsymbol{z}, 1/A) = c_r(\boldsymbol{z}, 1)/A$ and (39) follows. In words, if $g(c) = \log(c)$, then A can be computed purely from the growth in real consumption in any location in the decentralized equilibrium: $A = c_r(\boldsymbol{z}, 1)/c_r(\boldsymbol{z}^0, 1)$ — one does not even need to solve a counterfactual equilibrium with different factor-augmenting productivity.

One may wonder: why not always use the growth in real consumption to measure the growth in aggregate welfare? The reason is that outside the cases $g(c) = \log(c)$ and $\epsilon_r = \epsilon$, the growth in real consumption in each location, $c_r(\boldsymbol{z}, 1)/c_r(\boldsymbol{z}^0, 1)$, can be a different number, even though all agents are homogeneous in terms of welfare.[38] Our measure of aggregate welfare, A, in contrast provides a single number.[39]

A special case is when consumers only have preferences over consumption and do not directly value the location choice. In this case, the requirement that $g(c) = \log(c)$ is automatically satisfied and we get the following immediate corollary.

Corollary 6 (Preferences over consumption only)**.** *Suppose $\succeq_h$ ranks allocations of consumption and location (c_h, l_h) only by the consumption value c_h. That is, $(c_h, l_h) \succ_h (c'_h, l'_h)$ if, and only if, $c_h > c'_h$. If agents have homogeneous skills, $a_{hr} = a_r$ for every r and h, then*

$$A = \frac{c_r(\boldsymbol{z}, 1)}{c_r(\boldsymbol{z}^0, 1)},$$

[38] To see this, note that indifference between locations implies that $g(c_r) - g(c_{r'}) = \epsilon_{r'} - \epsilon_r$ must always hold. Hence, $g(c_r(\boldsymbol{z}, 1)) - g(c_r(\boldsymbol{z}^0, 1)) = g(c_{r'}(\boldsymbol{z}, 1)) - g(c_{r'}(\boldsymbol{z}^0, 1))$. This does not, in general, imply equal growth rates for real consumption per capita across locations unless $g(c) = \log(c)$.

[39] If $g(c) = \log(c)$, then one can also calculate A using a specific cardinalization of utility. In particular, if we use $f_h(x) = \exp(x)$ so that $u_h(c, r) = c \exp(\epsilon_r)$, then A is equal to the ratio of utilities in the counterfactual decentralized equilibrium relative to the status quo. This is a common approach in spatial models with homogeneous preferences (see e.g. Allen and Arkolakis, 2014). This trick does not work if $g(c)$ is not logarithmic. Moreover, one cannot simply assume that $g(c)$ is logarithmic without loss of generality: if ϵ_r varies by location, then the shape of $g(c)$ is determined by ordinal properties of the underlying agent-level utility functions because it disciplines income and substitution effects.

for any location r that is nonempty in both equilibria.

We now provide the proof of Proposition 9.

Proof of Proposition 9. Define $u_h(\boldsymbol{z}, Z)$ to be the utility of household h in a decentralized equilibrium given productivity vector $\boldsymbol{z}$ and aggregate factor augmenting productivity Z. Because households have identical ordinal preferences, skills, and opportunity sets, they attain the same welfare change and require the same compensating transfer at any given wages and prices. Aggregate budget balance therefore implies that this common transfer is zero. Hence, to calculate A, we simply need to solve

$$u_h(\boldsymbol{z}, 1/A) = u_h(\boldsymbol{z}^0, 1).$$

For any location r such that $L_r(\boldsymbol{z}, 1/A) \times L_r(\boldsymbol{z}^0, 1) > 0$ we have that

$$g(c_r(\boldsymbol{z}, 1/A)) + \epsilon_r = g(c_r(\boldsymbol{z}^0, 1)) + \epsilon_r.$$

It follows that

$$c_r(\boldsymbol{z}, 1/A) = c_r(\boldsymbol{z}^0, 1).$$

This completes the first part of the proposition.

Consider the case where $g(c) = \log c$. We need to show that

$$c_r\big(\boldsymbol{z}, 1/A\big) = A^{-1}\, c_r\big(\boldsymbol{z}, 1\big).$$

Let factor-augmenting productivity Z change from 1 to $1/A$ while holding $\boldsymbol{z}$ fixed. Conjecture an equilibrium allocation that scales all effective factor inputs, outputs, and intermediate inputs by $1/A$, keeping prices and wages (per efficiency unit) unchanged. In this conjectured equilibrium, consumption in every location scales with $1/A$. With logarithmic g, if consumption in every location is scaled by a common factor, then agents do not have any incentive to switch locations. Hence, labor supply in each location remains unchanged:

$$L_r\big(\boldsymbol{z}, 1/A\big) = L_r\big(\boldsymbol{z}, 1\big).$$

Effective factor supply is ZL_r, so it scales by $1/A$. Constant returns to scale then make it feasible for all outputs and intermediate inputs to scale by the same amount while relative prices remain unchanged. This confirms the conjectured equilibrium and completes the proof. □

A.3 Relation to the social surplus function

Suppose that each agent's preferences are ordinally additively separable, so that they can be represented by the utility function

$$u_h(c_r, r) = g(c_r) + \epsilon_{hr},$$

where $g(\cdot)$ is some increasing function and ϵ_{hr} are household-by-location specific taste shifters. The utilitarian social welfare function, given this cardinalization of utility and a vector of real wages across locations $\boldsymbol{w}/\boldsymbol{p}^c$, is

$$U(\boldsymbol{w}/\boldsymbol{p}^c) = \mathbb{E}[\max_{l_h} g(a_{hl_h} w_{l_h}/p^c_{l_h}) + \epsilon_{h,l_h}],$$

where $a_{hr}w_r/p^c_r$ is real consumption of agent h if location r is chosen. This social welfare function is called the "*Social Surplus Function*" in the discrete choice literature (see McFadden, 1981).[40]

The proposition below shows that under the assumptions of Proposition 3, the social surplus function can be used to calculate $A(t)$.

Proposition 10 (Using Social Surplus to Calculate A)**.** *Suppose the conditions of Proposition 3 hold. Then, TFP-equivalent aggregate surplus, A, satisfies the following equation:*

$$U(\boldsymbol{w}/(A\boldsymbol{p}^c)) = U(\boldsymbol{w}^0/\boldsymbol{p}^{c,0}),$$

where $\boldsymbol{w}/(A\boldsymbol{p}^c)$ are real wages in the decentralized equilibrium with productivities $\boldsymbol{z}$ and aggregate factor-augmenting productivity $1/A$ and $\boldsymbol{w}^0/\boldsymbol{p}^{c,0}$ are real wages in the status quo equilibrium. If, in addition, real wages do not depend on labor supplied in each region (e.g. as in the one-good economy), then we also have that $A = A^U$, where A^U is the consumption-equivalent variation defined in Equation (33) *for the social surplus function $W = U$.*

The proof is at the end of this subsection. In words, A can be computed as the reduction in labor productivity in every location such that the social surplus function, with productivity shocks $\boldsymbol{z}$, is equal to social surplus under the status quo (without the need to explicitly compute any compensating transfers). The assumptions of Proposition 3 are critical; if they do not hold, then the social surplus function U cannot be generally used to compute A (see Example 11 below).

[40] The social surplus function plays an important role in the literature in part because of the Williams-Daly-Zachary theorem (Williams, 1977; Daly and Zachary, 1978), which gives conditions under which the social surplus function acts like a potential function (i.e. its derivatives are equal to aggregate location choices).

Notably, even if the assumptions of Proposition 3 hold, A may still differ from the consumption-equivalent under the social surplus function, A^U. This is because real wages with productivities $(\boldsymbol{z}, 1/A)$ are not the same as real wages with productivities $\boldsymbol{z}$ divided by A. Intuitively, a reduction in aggregate factor productivity can cause agents to switch locations, and if real wages are a function of labor supply, then this can cause real wages to change and A to be different than A^U. The last part of the proposition points out that if, in addition to the assumptions of Proposition 3, real wages do not respond to changes in labor supplies, as in the one-good economy where they are pinned down by productivity parameters, then A coincides with A^U.

Proposition 10 relates to a well-known property of discrete choice models where consumers choose among a discrete set of options given fixed prices and incomes. In that literature, it is known that if the indirect utility function is linear in the price of the good, then the social surplus function can be used to calculate the sum of compensating variations (Small and Rosen, 1981).[41]

The example below illustrates Proposition 10 using a simple example.

Example 10 (Using Social Surplus to Calculate Aggregate Welfare). Consider the one-good economy example from Example 3 again. There is a single freely-traded consumption good produced linearly from labor and productivity in each location r is z_r. We impose the assumptions of Proposition 3 and Assumption 1. The real wage per efficiency unit in each location is simply $w_r = z_r$. Hence, by Proposition 10, we have

$$U(\boldsymbol{z}/A) = \mathbb{E}[\max_r \frac{z_r}{A} + \epsilon_{hr}] = \mathbb{E}[\max_r z_r^0 + \epsilon_{hr}] = U(\boldsymbol{z}^0).$$

If the ϵ_{hr} are independently drawn from Type I extreme value distributions (i.e., Gumbel), then

$$U(\boldsymbol{z}/A) = \frac{1}{\theta} \log \left[\sum_r \exp\left(\theta z_r/A + B_r\right) \right] + \text{constant}.$$

[41] In the absence of linearity, calculating the sum of compensating variations typically requires resorting to simulation methods. See Hortaçsu and Joo (2023), and the references therein, for a recent textbook discussion. As discussed by Hortaçsu and Joo (2023), many recent studies in the industrial organization literature ignore individual compensating variations and directly use the social surplus function as the starting point of their analysis, even if the quasi-linear justification for this approach does not hold. In this case, the social surplus function is interpreted like a social welfare function. Dubé et al. (2025) analyze the welfare properties of this social surplus function for a broad class of demand systems.

Therefore, since the logarithm is strictly increasing, $U(\boldsymbol{z}/A) = U(\boldsymbol{z}^0)$ if and only if

$$\sum_r \exp\left(\theta z_r/A + B_r\right) = \sum_r \exp\left(\theta z_r^0 + B_r\right),$$

which is equation (18). Furthermore, in this one-good economy, real wages in each location are equal to productivity and independent of labor supply. Hence, scaling all productivities by $1/A$ is observationally equivalent to dividing all real wages by A. This makes $A = A^U$.

Proposition 10 does not generally extend when the assumptions of Proposition 3 do not hold: i.e. if either $g(c)$ is nonlinear or if consumption prices vary by location. Intuitively, when the assumptions of Proposition 3 hold, we have a transferable utility model where the sum of $g(c_h)$ across households is interpretable as an aggregate resource constraint and can be split across households freely using lump-sum transfers. However, if either $g(c)$ is nonlinear or agents in different locations consume different consumption goods, then utility is not transferable and the sum of $g(c_h)$ across households cannot be interpreted as an aggregate resource constraint.

The following example demonstrates that when the assumptions of Proposition 3 do not hold, as in the commonly studied Fréchet case, then A^U and A need not agree even to a first order approximation (and can even have different signs in logs).

Example 11 (When Proposition 10 Fails). Consider Example 5 again. There is a common freely-traded consumption good produced linearly from labor, with productivity z_r in location r. The assumptions of Proposition 3 do not hold in this example because $g(c)$ is logarithmic. Assume that ϵ is distributed according to a Gumbel distribution with inverse scale θ, and workers have homogeneous skills. Utilitarian welfare, as measured by A^U, is the solution to

$$U(\boldsymbol{z}/A^U) = \mathbb{E}[\max_r \log(z_r/A^U) + \epsilon_{hr}] = \mathbb{E}[\max_r \log(z_r^0) + \epsilon_{hr}] = U(\boldsymbol{z}^0).$$

One can show that $A^U = \left(\sum_r (B_r z_r)^\theta / \sum_r (B_r z_r^0)^\theta\right)^{\frac{1}{\theta}}$.[42] To a first-order approximation,

[42] Compare this with the right hand side of (34), which calculates A^W under a different cardinalization, $\exp(\log(c_r) + \epsilon_{hr})$. The fact that the consumption-equivalent variation of the social welfare function $U(\boldsymbol{c}) = \mathbb{E}[\max_{l_h} \log(c_{l_h}) + \epsilon_{hl_h}]$ is equal to that of the social welfare function $W(\boldsymbol{c}) = \mathbb{E}[\max_{l_h} c_{l_h} \exp(\epsilon_{hl_h})]$

changes in $\log A^U$ are a population-weighted sum of changes in productivities:

$$\Delta \log A^U \approx \sum_r L_r^0 \Delta \log z_r,$$

where L_r^0 is the share of the population choosing location r at the status quo. This does not coincide, even to a first-order, with changes in aggregate welfare as measured by A, because $\Delta \log A \approx \sum_r \lambda_r^0 \Delta \log z_r$, where the Domar weights in this example are $\lambda_r^0 = L_r^0 z_r^0 / (\sum_{r'} L_{r'}^0 z_{r'}^0)$. It can easily be the case that $\Delta \log A > 0$, so that winners can compensate losers and have resources left over, and yet $\Delta \log A^U < 0$.

Proof of Proposition 10. Without loss of generality, let the common consumption good price p^c be the numeraire. The social surplus function is

$$U(\boldsymbol{w}) = \mathbb{E}_h \left[\max_r \{ a_{hr} w_r + \epsilon_{hr} \} \right],$$

where w_r is now the real wage in location r. Then,

$$\frac{\partial U}{\partial w_r} = \mathbb{E}_h \left[a_{hr} \mathbf{1} \left\{ r \in \arg\max_j (a_{hj} w_j + \epsilon_{hj}) \right\} \right] = L_r(\boldsymbol{w}).$$

Equation (43) in the proof of Theorem 2 states that with a common consumption good

$$\sum_r L_r^{\text{comp}} \left[\frac{w_r}{A} \right] d \log \left[\frac{w_r}{A} \right] = \sum_r L_r^{\text{comp}} \left[\frac{w_r}{A} \right] d \log p^c,$$

where the wages and prices are the ones in the compensated equilibrium. Since the consumer price index is the numeraire, this is the same as:

$$\sum_r L_r^{\text{comp}} \left[\frac{w_r}{A} \right] d \log \left[\frac{w_r}{A} \right] = 0.$$

Equivalently,

$$\sum_r L_r^{\text{comp}} d \left[\frac{w_r}{A} \right] = 0.$$

Index any continuous path between the status quo and the counterfactual allocation by s

is a special implication from the assumption that ϵ is distributed according to a Gumbel distribution. See Proposition 11, below, for more details.

between 0 and 1. The fundamental theorem of calculus gives:

$$\int_{\boldsymbol{w}(0)}^{\boldsymbol{w}(1)/A} \sum_r L_r^{\text{comp}}(\boldsymbol{x})\, dx_r = 0.$$

By Proposition 3, the compensated and uncompensated labor supply functions are the same. Furthermore, given productivities $\boldsymbol{z}$ and aggregate factor-augmenting productivity $1/A$, the compensated and uncompensated equilibrium real wages are the same. Hence, we can write

$$\int_{\boldsymbol{w}(0)}^{\boldsymbol{w}(1)/A} \sum_r L_r(\boldsymbol{x})\, dx_r = 0,$$

substituting in $\frac{\partial U}{\partial w_r}$ from above gives

$$\int_{\boldsymbol{w}(0)}^{\boldsymbol{w}(1)/A} \sum_r \frac{\partial U}{\partial w_r}(\boldsymbol{x})\, dx_r = 0,$$

which is equal to

$$U(\boldsymbol{w}(1)/A) - U(\boldsymbol{w}(0)) = 0,$$

by the fundamental theorem of calculus. Finally, if real wages do not depend on labor supply, real wages in the compensated equilibrium coincide with those in the decentralized equilibrium at t. The preceding equation is then precisely the equation defining A^U, so $A = A^U$. □

A.3.1 Invariance of A^W under common cardinal transformations

The specification with ordinally additive preferences and Gumbel-distributed tastes has an invariance property, stated in Proposition 11 below. Specifically, in this case A^W is invariant to the choice of monotone transformations as long as the transformation is common for all agents. This invariance, however, requires arbitrarily restricting the set of allowable monotone transformations of utility. For example, under the restrictions in Proposition 11, differences in the overall level of tastes across households are left unnormalized: some households have high values of ϵ_{hr} in every location, while others have low values in every location. The cardinalizations we discuss in Section 6.3.2, for example, fall outside this restricted class.

Proposition 11 (Invariance to Restricted Monotone Transformations)**.** *Suppose that*

$$u_h(c_r, r) = f_h\big(g(c_r) + \epsilon_{hr}\big),$$

where $f_h = f$ is common across households, and the ϵ_{hr} are independent Gumbel random variables with inverse scale $\theta > 0$ and location parameters $\log B_r$ with $B_r > 0$. Then A^W is the unique solution to:

$$\sum_r B_r^\theta \exp\left[\theta g\left(\frac{c_r(t)}{A^W(t)}\right)\right] = \sum_r B_r^\theta \exp\left[\theta g(c_r(0))\right].$$

Hence, A^W does not depend on f. If, in addition, $g(c) = \log c$, then

$$A^W = \left[\frac{\sum_r (B_r c_r)^\theta}{\sum_r (B_r c_r^0)^\theta}\right]^{1/\theta}.$$

Proof. Since f is strictly increasing,

$$\max_r u_h(c_r, r) = f\left(\max_r\{g(c_r) + \epsilon_{hr}\}\right).$$

For any $v \in \mathbb{R}$, independence across locations gives the following CDF for the maximum

$$\begin{aligned}
\Pr\left(\max_r\{g(c_r) + \epsilon_{hr}\} \le v\right) &= \prod_r \Pr\left(\epsilon_{hr} \le v - g(c_r)\right) \\
&= \exp\left\{-\left[\exp(-\theta v)\right]\sum_r B_r^\theta \exp\left[\theta g(c_r)\right]\right\} \\
&= \exp\left\{-\exp\left[-\theta\left(v - \frac{1}{\theta}\log\left(\sum_r B_r^\theta \exp\left[\theta g(c_r)\right]\right)\right)\right]\right\}.
\end{aligned}$$

Thus, the maximum is Gumbel with inverse scale θ and location parameter

$$\frac{1}{\theta}\log\left(\sum_r B_r^\theta \exp\left[\theta g(c_r)\right]\right).$$

For a fixed θ, the consumption vector $\boldsymbol{c}$ affects the distribution of the maximum only through this location parameter. Since f is common and strictly increasing, average maximized utility is strictly increasing in the location parameter. The definition of A^W therefore implies

$$\frac{1}{\theta}\log\left(\sum_r B_r^\theta \exp\left[\theta g\left(\frac{c_r}{A^W}\right)\right]\right) = \frac{1}{\theta}\log\left(\sum_r B_r^\theta \exp\left[\theta g(c_r^0)\right]\right).$$

Dropping the log, this is the equation in the statement of the proposition. The left-hand side is decreasing in A^W because g is increasing. Hence the solution is unique. Because the equation contains no f, neither does A^W. Finally, if $g(c) = \log c$, the equation becomes

$$(A^W)^{-\theta} \sum_r (B_r c_r)^\theta = \sum_r (B_r c_r^0)^\theta.$$

□

B Appendix for Section 4

B.1 Input-Output Notation for this Appendix

We introduce some useful accounting variables. Let λ be the vector of Domar (1961) weights, whose ith component is commodity i's sales or factor i's income divided by total income:

$$\lambda_i = \frac{p_i y_i}{\sum_r Z w_r L_r} \mathbf{1}[i \in C + N] + \frac{Z w_i L_i}{\sum_r Z w_r L_r} \mathbf{1}[i \in R].$$

Let Ω be the $|C + N + R| \times |C + N + R|$ input-output matrix. The first $|C|$ rows and columns are the consumption bundles, the next $|N|$ are the intermediates, and the final $|R|$ are the primary factors. For $i \in C + N$, the ijth element of Ω is the expenditure share of i on inputs from j relative to i's total sales:

$$\Omega_{ij} = \frac{p_j x_{ij}}{p_i y_i} \mathbf{1}[j \in N] + \frac{w_j l_{ij}}{p_i y_i} \mathbf{1}[j \in R].$$

The rows corresponding to primary factors are zero.

Define the spending shares to be:

$$\chi_r = \frac{E_r}{\sum_s E_s}.$$

Market clearing for $i \in C + N + R$ requires that commodity i's sales, or factor i's income, equal total spending on i by households and producers:

$$\lambda_i = \sum_{r \in R} \chi_r \mathbf{1}[i = c(r)] + \sum_{j \in C+N} \lambda_j \Omega_{ji}, \tag{40}$$

where $c(r)$ is the consumption bundle associated with location r, and χ_r is the share of final consumption spending in location r. The factor market clearing condition states that

spending on factor r by firms is equal to earnings by agents that supply that factor:

$$\lambda_r = \frac{Z w_r L_r}{\sum_{r'} Z w_{r'} L_{r'}}. \tag{41}$$

The market clearing conditions in the compensated equilibrium are the same as (40) and (41) except where we use compensated supply L_r^{comp} and compensated demand $\chi_r^{\text{comp}}(\boldsymbol{w}/A, \boldsymbol{p}^c) = \frac{E_r^{\text{comp}}(\boldsymbol{w}/A, \boldsymbol{p}^c)}{\sum_s E_s^{\text{comp}}(\boldsymbol{w}/A, \boldsymbol{p}^c)}$ in place of the uncompensated supply and demand functions.

B.2 Proofs for Section 4

Proof of Theorem 2. Throughout the proof, wages, prices, quantities, and expenditure shares refer to the compensated equilibrium, so we suppress the superscript comp. Budget balance in the compensated equilibrium requires

$$\int_h T_h(\boldsymbol{w}/A, \boldsymbol{p}^c)\, dh = 0. \tag{42}$$

Differentiating (42) with respect to $\boldsymbol{w}/A$, and $\boldsymbol{p^c}$ and applying Proposition 2 gives

$$0 = -\sum_{r \in R} L_r^{\text{comp}} d(w_r/A) + \sum_{r \in R} \frac{E_r^{\text{comp}}}{p_r^c} dp_r^c,$$

and hence

$$\sum_{r \in R} (w_r/A) L_r^{\text{comp}} d\log(w_r/A) = \sum_{r \in R} E_r^{\text{comp}} d\log p_r^c. \tag{43}$$

Because aggregate transfers add up to zero,

$$\sum_{r \in R} E_r^{\text{comp}} = \sum_{r \in R} (w_r/A) L_r^{\text{comp}}.$$

Dividing (43) by this common value and using $d\log(w_r/A) = d\log w_r - d\log A$ gives

$$d\log A = \sum_{r \in R} \lambda_r d\log w_r - \sum_{r \in R} \chi_r d\log p_r^c, \tag{44}$$

where $\lambda_r = w_r L_r^{\text{comp}} / (\sum_{r'} w_{r'} L_{r'}^{\text{comp}})$ and $\chi_r = E_r^{\text{comp}} / (\sum_{r'} E_{r'}^{\text{comp}})$. We now connect wages and consumption prices to productivities. For every produced commodity $i \in C + N$,

Shephard's lemma and perfect competition imply

$$d\log p_i = -d\log z_i + \sum_{j\in N}\Omega_{ij}d\log p_j + \sum_{r\in R}\Omega_{ir}d\log w_r.$$

Multiplying by λ_i and summing over $i \in C+N$ gives

$$\begin{aligned}\sum_{i\in C+N}\lambda_i d\log p_i = &-\sum_{i\in C+N}\lambda_i d\log z_i \\ &+\sum_{j\in N}\left(\sum_{i\in C+N}\lambda_i\Omega_{ij}\right)d\log p_j \\ &+\sum_{r\in R}\left(\sum_{i\in C+N}\lambda_i\Omega_{ir}\right)d\log w_r.\end{aligned}$$

Market clearing implies that the two expressions in parentheses are λ_j and λ_r, respectively. Moreover, $\lambda_{c(r)} = \chi_r$ and $p_{c(r)} = p_r^c$. Separating the consumption bundles from the intermediate goods on the left-hand side and cancelling the intermediate-price terms therefore gives

$$\sum_{r\in R}\chi_r d\log p_r^c = -\sum_i \lambda_i d\log z_i + \sum_{r\in R}\lambda_r d\log w_r.$$

Substitution into (44) yields

$$d\log A = \sum_{i\in C+N}\lambda_i d\log z_i.$$

Bringing back the compensated-equilibrium superscript and integrating along the path from 0 to 1 gives

$$\log A = \int_0^1 \sum_{i\in C+N}\lambda_i^{\text{comp}}(s)\frac{d\log z_i(s)}{ds}\,ds, \tag{45}$$

completing the proof. □

B.3 Second-Order Approximation

Corollary 4 shows that approximating A to a second order requires knowledge of $d\lambda^{\text{comp}}$. Here, we show how to solve for this using the same approach as in Baqaee and Farhi (2019). For simplicity of exposition, assume that all production and consumption functions are CES, with the ith producer in the input-output matrix Ω corresponding to a CES aggregator with elasticity of substitution θ_i. This structure is flexible enough to capture

any nested-CES production and consumption functions through relabeling.[43] For additional simplicity, we also assume that households have homogeneous skills $a_{hr} = 1$ for every h and r (one can extend the results by simply conditioning labor supply by skill). We also assume that transfers in the decentralized equilibrium are zero.

Under these simplifying assumptions, the following proposition determines sales shares in both the compensated and uncompensated equilibrium.

Proposition 12 (Prices and Expenditures in Equilibrium)**.** *In the decentralized equilibrium, given labor supplies $\boldsymbol{L}$ and final spending shares $\boldsymbol{\chi}$, the following differential equations determine the rest of the equilibrium variables:*

$$d\log\Omega_{ij} = (1-\theta_i)\left[d\log p_j \mathbf{1}[j\in N] + d\log w_j \mathbf{1}[j\in R] - \sum_{k\in N}\Omega_{ik} d\log p_k - \sum_{r\in R}\Omega_{ir} d\log w_r\right], \quad (\Omega_{ij}\neq 0) \tag{46}$$

$$d\log p_i = -d\log z_i + \sum_{j\in N}\Omega_{ij} d\log p_j + \sum_{r\in R}\Omega_{ir} d\log w_r, \quad (i\in C+N) \tag{47}$$

$$d\log\lambda_r = d\log w_r + d\log L_r - \sum_{r'}\lambda_{r'}\left[d\log w_{r'} + d\log L_{r'}\right], \quad (r\in R) \tag{48}$$

$$d\log\lambda_i = \lambda_i^{-1}\left(\sum_{r\in R} d\chi_r \mathbf{1}[i=c(r)] + \sum_{j\in C+N}\left[d\lambda_j\Omega_{ji} + \lambda_j d\Omega_{ji}\right]\right), \tag{49}$$

assuming the equilibrium system is nonsingular. If we swap $\boldsymbol{L}$ and $\boldsymbol{\chi}$ for $\boldsymbol{L}^{comp}$ and $\boldsymbol{\chi}^{comp}$, we get the evolution of variables in the compensated equilibrium instead. The boundary conditions are the same for both the compensated and decentralized equilibria — spending shares are equal to their values in status quo, and initial prices and wages are determined up to a common numeraire.

Proposition 12 can be used to calculate either the decentralized equilibrium, given $d\log\chi_r$ and $d\log L_r$, or the compensated equilibrium, given $d\log\chi_r^{\text{comp}}$ and $d\log L_r^{\text{comp}}$. The changes $d\log\chi_r^{\text{comp}}$ and $d\log L_r^{\text{comp}}$ can be computed numerically using Proposition 1. This approach is useful both computationally and analytically. Computationally, it obviates the need to solve a system of nonlinear equations to compute $A(t)$ (i.e. as in Theorem 1). Instead, it requires repeatedly solving a system of linear equations and updating variables (Appendix B.4 gives the numerical procedure). Analytically, it allows us to prove some useful results.

Before using Proposition 12, we briefly describe the intuition for equations (46)-(49). The expression in (46) is the log-linearized input expenditure share of i for input j, which depends on the elasticity of substitution, θ_i, and on how the price of j changes relative to

[43] One can extend to non-nested CES along similar lines to Baqaee and Farhi (2019).

the marginal cost of i. The expression in (47) is Shephard's lemma, relating the marginal cost of i to the expenditure-share weighted change in input prices and the productivity shock to i. Equation (48) log-linearizes the definition of λ_r, and (49) gives the log-linearized market-clearing conditions (40).

Proof of Proposition 12. Equation (46) follows by log-differentiating the CES input-demand equations. Equation (47) follows from Shephard's lemma, constant returns to scale, and the condition that price equals marginal cost. Equation (48) follows by log-differentiating the definition of λ_r for $r \in R$, and (49) follows by differentiating the goods-market-clearing condition (40). These cost-minimization, pricing, and market-clearing conditions hold in both the decentralized and compensated equilibria. □

Calculating $\Delta\lambda_i^{\text{comp}}$ in Corollary 4 is a straightforward application of Proposition 12 if we know the derivatives of compensated labor supply dL^{comp} and demand $d\chi^{\text{comp}}$ in the status quo. When the assumptions of Proposition 3 hold, this is simple to do using Proposition 3, since dL^{comp} can be computed using the uncompensated supply system and $d\chi^{\text{comp}}$ is irrelevant. In this subsection, we focus on characterizing dL^{comp} and $d\chi^{\text{comp}}$ at the status quo, without imposing the assumptions of Proposition 3, and we compare these to their decentralized uncompensated counterparts.

We now proceed to characterizing $d\chi^{\text{comp}}$ (in Proposition 13) and dL^{comp} (in Proposition 14) at the status quo. Given these derivatives, we can compute $d\lambda^{\text{comp}}$ using Proposition 12, and hence, we can approximate $\log A(t)$ to a second-order using Corollary 4.

Proposition 13 (First-Order Change in Spending Shares)**.** *For any variable that varies by location, x_r, define $\mathbb{E}_\chi[x] \equiv \sum_r \chi_r(0)x_r$. Starting at the status quo, the change in spending share in location r, in the decentralized equilibrium, satisfies*

$$d\log\chi_r = d\log w_r - \mathbb{E}_\chi\left[d\log w\right] + d\log L_r - \mathbb{E}_\chi\left[d\log L\right],$$

where the changes in wages $\boldsymbol{w}$ and locations $\boldsymbol{L}$ are in the decentralized equilibrium. In the compensated equilibrium, we instead have

$$d\log\chi_r^{comp} = d\log p_r^c - \mathbb{E}_\chi\left[d\log p^c\right] + d\log L_r^{comp} - \mathbb{E}_\chi\left[d\log L^{comp}\right],$$

where the changes in prices $\boldsymbol{p}^c$ and locations $\boldsymbol{L}^{comp}$ are in the compensated equilibrium.

The intuition is the following. In the decentralized equilibrium, the share of spending by each location must move in line with the share of income earned by households in that location. This is the product of wages w_r and population shares L_r. A location's

share of spending rises if wages in that location increase quickly or if households flock to that location. On the other hand, in the compensated equilibrium, the share of spending in each location depends on the rate at which the cost of consumption in location r has increased and the population shares in the compensated equilibrium. If consumption prices rise quickly in a location, or households flock to a location, in the compensated equilibrium, then that location's share of total spending rises.

Proof of Proposition 13. In the decentralized equilibrium, transfers are zero, so expenditure in location r is

$$E_r = w_r L_r.$$

Log-differentiating $\chi_r = E_r / \sum_{r'} E_{r'}$ at $t = 0$ therefore gives

$$d \log \chi_r = d \log w_r + d \log L_r - \mathbb{E}_\chi \left[d \log w + d \log L \right].$$

Now consider the compensated equilibrium. Index a path from $\boldsymbol{z}^0$ to $\boldsymbol{z}$ by t and let $B_r^{\text{comp}}(t)$ denote the set of households whose compensated choice at t is r. Expenditure in location r is

$$E_r^{\text{comp}}(t) = \int_{B_r^{\text{comp}}(t)} p_r^c(t) \bar{c}_{hr} \, dh.$$

By definition, $\bar{c}_{hr}$ is the consumption required in location r to make the household indifferent to the status quo, so it does not vary with t. At the status quo, $p_r^c \bar{c}_{hr} = w_r$ for all households in $B_r^{\text{comp}}(0)$ and on its boundary. Assumption 1 implies that

$$L_r^{\text{comp}}(t) = \int_{B_r^{\text{comp}}(t)} dh.$$

Under the regularity conditions ensuring differentiability of the choice shares, applying Leibniz's rule to the moving choice set therefore gives

$$dE_r^{\text{comp}} = w_r L_r \, d \log p_r^c + w_r \, dL_r^{\text{comp}},$$

where $d \log p_r^c$ is the change in consumption prices in the compensated equilibrium. The first term is the change in expenditure by households initially choosing r, while the second is the contribution from households entering or leaving r. Using

$$E_r^{\text{comp}}(0) = w_r L_r \qquad \text{and} \qquad L_r^{\text{comp}}(0) = L_r,$$

it follows that

$$d \log E_r^{\text{comp}} = d \log p_r^c + d \log L_r^{\text{comp}}.$$

Finally, log-differentiating

$$\chi_r^{\text{comp}} = \frac{E_r^{\text{comp}}}{\sum_{r'} E_{r'}^{\text{comp}}}$$

gives

$$d \log \chi_r^{\text{comp}} = d \log p_r^c + d \log L_r^{\text{comp}} - \sum_{r'} \chi_{r'}^{\text{comp}}(0) \left[d \log p_{r'}^c + d \log L_{r'}^{\text{comp}} \right].$$

Since the compensated and decentralized equilibria coincide at the status quo, $\chi^{\text{comp}}(0) = \chi(0)$. Hence,

$$d \log \chi_r^{\text{comp}} = d \log p_r^c + d \log L_r^{\text{comp}} - \mathbb{E}_\chi \left[d \log p^c + d \log L^{\text{comp}} \right],$$

as required. □

The next proposition provides a general characterization of dL^{comp} in terms of the derivatives of the uncompensated labor supply function at the status quo $\partial L_r / \partial (w_{r'} / p_{r'}^c)$. This is important because the derivatives of the uncompensated supply system are estimable using exogenous variation in the decentralized equilibrium.[44]

For the characterization of uncompensated flows, we assume that preferences are *ordinally additively separable* between location choices and consumption. This means that the utility function can be written as

$$u_h(c_h, l_h) = f_h(g(c_h) + \epsilon_{h l_h}) \tag{50}$$

where g is a strictly increasing function, $\epsilon_{h l_h}$ is a taste parameter for household h's preferences for location l_h, and f_h is any strictly increasing (potentially household-specific) function.[45]

Proposition 14 (First-Order Changes in Location Choices)**.** *The change in the share of households in location r can be decomposed into inflows minus outflows:*

$$dL_r = \sum_{r' \neq r} dL_{r' \to r} - \sum_{r' \neq r} dL_{r \to r'}, \tag{51}$$

[44] With zero transfers, consumption in location r equals w_r / p_r^c, so uncompensated labor supply is $L_r(\boldsymbol{w}/\boldsymbol{p}^c)$. This allows us to define $\partial L_r / \partial (w_{r'} / p_{r'}^c)$, which can be estimated using the observed changes in market-level labor supply in response to exogenous variation in real wages.

[45] Since utility is only pinned down up to monotone increasing transformations, the specific function f_h has no testable implications and can vary at the level of individual agents in arbitrary ways so long as it is increasing. In contrast to f_h, the shape of the function g does have testable implications and its shape can be recovered from choice data. See Allen and Rehbeck (2019) for more information on how preferences can be identified from behavior in discrete choice settings.

where $dL_{r' \to r}$ is the share of households switching from r' to r. Under ordinal additive separability, as in (50)*, these flows satisfy, to a first order:*

$$dL_{r \to r'} = \left[\frac{\partial L_{r'}}{\partial (w_r / p_r^c)} d(\frac{w_r}{p_r^c}) - \frac{\partial L_r}{\partial (w_{r'} / p_{r'}^c)} d(\frac{w_{r'}}{p_{r'}^c}) \right] \mathbf{1} \left[\frac{\partial L_r}{\partial (w_{r'} / p_{r'}^c)} d(\frac{w_{r'}}{p_{r'}^c}) \leq \frac{\partial L_{r'}}{\partial (w_r / p_r^c)} d(\frac{w_r}{p_r^c}) \right], \tag{52}$$

where wages and prices are those of the decentralized equilibrium.

The compensated supply function instead satisfies:

$$dL_r^{comp} = \sum_{r' \neq r} dL_{r' \to r}^{comp} - \sum_{r' \neq r} dL_{r \to r'}^{comp},$$

where $dL_{r' \to r}^{comp}$ is the share of households switching from r' to r in the compensated equilibrium. At $t = 0$, these flows satisfy:

$$dL_{r \to r'}^{comp} = \frac{\partial L_r}{\partial \left[w_{r'} / p_{r'}^c \right]} \frac{1}{p_{r'}^c} \left(p_r^c d(\frac{w_r}{A p_r^c}) - p_{r'}^c d(\frac{w_{r'}}{A p_{r'}^c}) \right) \times \mathbf{1} \left[p_r^c d(\frac{w_r}{A p_r^c}) \leq p_{r'}^c d(\frac{w_{r'}}{A p_{r'}^c}) \right], \tag{53}$$

where wages and prices are those of the compensated equilibrium and, at $t = 0$, $d \log A$ is given by Corollary 2.

The key equation of Proposition 14 is (53), which expresses compensated flows in terms of uncompensated labor supply derivatives. This equation does not require ordinal additive separability. We now discuss flows in the decentralized and compensated equilibria in turn.

In the decentralized equilibrium, Equation (52) expresses changes in location choices as pairwise movements across indifference boundaries. Changes in real wages determine the direction in which the boundary between r and r' moves, while the cross-derivatives of labor supply between r and r' determine the mass of households that crosses it. The overall change in the share of households choosing r, given by (51), is the sum of inflows into r minus outflows from r.

Equation (53) is the counterpart to (52) in the compensated equilibrium. Household-specific transfers keep every household indifferent to the status quo, while aggregate factor productivity is adjusted so that these transfers satisfy budget balance. For a household initially indifferent between r and r', the term $p_r^c d \left(\frac{w_r}{A p_r^c} \right)$ measures the first-order reduction in the transfer required for indifference if the household chooses r. The household chooses the location for which this reduction is larger. The size of the resulting flow is determined by the cross-derivative of labor supply between the two locations. Crucially, the partial derivatives in (53) are those of the uncompensated supply system evaluated at

the status quo in the decentralized economy.

Proof of Proposition 14. We first derive the expression for pairwise flows in the uncompensated equilibrium and relate them to derivatives of the uncompensated supply system. We then derive the compensated flows and express them in terms of those derivatives.

Uncompensated Equilibrium. Fix two distinct locations i and j. Under the regularity conditions ensuring differentiability of the choice shares, households close to more than one pairwise choice boundary do not affect the first-order calculation. We can therefore study the flow across the i-j boundary in isolation. Suppressing the household index, consider the portion of the status-quo choice boundary where households are indifferent between i and j and strictly prefer these two locations to every other location. With ordinal additive separability,

$$g\left(\frac{w_i}{p_i^c}\right) + \epsilon_i = g\left(\frac{w_j}{p_j^c}\right) + \epsilon_j > g\left(\frac{w_k}{p_k^c}\right) + \epsilon_k, \qquad k \notin \{i, j\}.$$

The cutoff for $\epsilon_j - \epsilon_i$ is

$$g\left(\frac{w_i}{p_i^c}\right) - g\left(\frac{w_j}{p_j^c}\right).$$

Equality corresponds to indifference between i and j. Above the cutoff, j is strictly preferred to i; below the cutoff, i is strictly preferred to j. Let κ_{ij} denote the density, measured with respect to $\epsilon_j - \epsilon_i$, of households on this portion of the boundary.

Let t index a path away from the status quo, attained at $t = 0$. Throughout the proof, d denotes the right derivative at $t = 0$, and undated variables are evaluated at the status quo. The change in the cutoff is

$$g'\left(\frac{w_i}{p_i^c}\right) d\left(\frac{w_i}{p_i^c}\right) - g'\left(\frac{w_j}{p_j^c}\right) d\left(\frac{w_j}{p_j^c}\right).$$

If this expression is negative, the cutoff falls, and the mass moving from i to j equals κ_{ij} times the fall in the cutoff. Hence,

$$\begin{aligned} dL_{i\to j} = \kappa_{ij} &\left[g'\left(\frac{w_j}{p_j^c}\right) d\left(\frac{w_j}{p_j^c}\right) - g'\left(\frac{w_i}{p_i^c}\right) d\left(\frac{w_i}{p_i^c}\right)\right] \\ &\times \mathbf{1}\left[g'\left(\frac{w_i}{p_i^c}\right) d\left(\frac{w_i}{p_i^c}\right) \le g'\left(\frac{w_j}{p_j^c}\right) d\left(\frac{w_j}{p_j^c}\right)\right]. \end{aligned} \tag{54}$$

We now express κ_{ij} in terms of cross-derivatives of the uncompensated supply system. All such derivatives are evaluated at the decentralized status quo. If w_j/p_j^c increases while all other real wages remain fixed, the cutoff falls and households leave i for j. Therefore,

$$\frac{\partial L_i}{\partial\left(w_j/p_j^c\right)} = -\kappa_{ij} g'\left(\frac{w_j}{p_j^c}\right).$$

Similarly, if w_i/p_i^c increases while all other real wages remain fixed, households leave j for i, so

$$\frac{\partial L_j}{\partial\left(w_i/p_i^c\right)} = -\kappa_{ij} g'\left(\frac{w_i}{p_i^c}\right).$$

Differentiability makes these equations valid for changes of either sign. Substituting them into (54) gives

$$\begin{aligned} dL_{i\to j} = & \left[\frac{\partial L_j}{\partial\left(w_i/p_i^c\right)} d\left(\frac{w_i}{p_i^c}\right) - \frac{\partial L_i}{\partial\left(w_j/p_j^c\right)} d\left(\frac{w_j}{p_j^c}\right)\right] \\ & \times \mathbf{1}\left[\frac{\partial L_i}{\partial\left(w_j/p_j^c\right)} d\left(\frac{w_j}{p_j^c}\right) \le \frac{\partial L_j}{\partial\left(w_i/p_i^c\right)} d\left(\frac{w_i}{p_i^c}\right)\right], \end{aligned}$$

which is equation (52). Summing pairwise inflows and subtracting pairwise outflows gives

$$dL_i = \sum_{j\neq i} dL_{j\to i} - \sum_{j\neq i} dL_{i\to j}.$$

Compensated Equilibrium. Now let changes in wages, prices, and A be those of the compensated equilibrium. Fix distinct locations i and j and consider households initially choosing i and close to indifference between i and j, with both locations preferred to all others. Let $T_i(t)$ keep these households indifferent to the status quo if they continue to choose i. Under Assumption 1, they have the same status-quo consumption, so

$$\frac{w_i(t)/A(t) + T_i(t)}{p_i^c(t)} = \frac{w_i^0}{p_i^{c,0}}.$$

Household h requires a smaller transfer for indifference in j than in i when

$$\frac{w_j(t)/A(t) + T_i(t)}{p_j^c(t)} > \bar{c}_{hj},$$

in which case j is its compensated choice by Proposition 1. For households initially choosing i, $\bar{c}_{hj} \geq w_j^0/p_j^{c,0}$, with equality for those initially indifferent between i and j.

Since $A(0) = 1$ and $T_i(0) = 0$, differentiating the condition that consumption in i remains fixed gives

$$dT_i = -p_i^c d\left(\frac{w_i}{Ap_i^c}\right).$$

Consumption in j under this transfer therefore changes by

$$d\left(\frac{w_j/A + T_i}{p_j^c}\right) = d\left(\frac{w_j}{Ap_j^c}\right) + \frac{dT_i}{p_j^c} = \frac{1}{p_j^c}\left[p_j^c d\left(\frac{w_j}{Ap_j^c}\right) - p_i^c d\left(\frac{w_i}{Ap_i^c}\right)\right]. \tag{55}$$

Consider an uncompensated increase in consumption in j, holding consumption elsewhere fixed. Every household leaving i goes to j, and no household enters i. Hence, the mass switching per unit increase in consumption in j is

$$-\frac{\partial L_i}{\partial(w_j/p_j^c)}.$$

Both this experiment and the compensated comparison under T_i keep consumption in i at its status-quo level. In either case, household h prefers j to i when the increase in consumption in j exceeds $\bar{c}_{hj} - w_j^0/p_j^{c,0}$. As in the uncompensated calculation, households close to indifference among three or more locations do not affect first-order flows. Changes in other locations therefore do not affect this pairwise calculation. The same switching rate therefore applies to the consumption change in (55), when that change is positive, giving

$$dL_{i\to j}^{\text{comp}} = \frac{\partial L_i}{\partial(w_j/p_j^c)}\frac{1}{p_j^c}\left[p_i^c d\left(\frac{w_i}{Ap_i^c}\right) - p_j^c d\left(\frac{w_j}{Ap_j^c}\right)\right] \times \mathbf{1}\left[p_i^c d\left(\frac{w_i}{Ap_i^c}\right) \leq p_j^c d\left(\frac{w_j}{Ap_j^c}\right)\right],$$

which is (53). The partial derivative is evaluated at the decentralized status quo, whereas changes in wages, prices, and A are those of the compensated equilibrium. Summing pairwise inflows and subtracting outflows gives

$$dL_i^{\text{comp}} = \sum_{j\neq i} dL_{j\to i}^{\text{comp}} - \sum_{j\neq i} dL_{i\to j}^{\text{comp}}.$$

□

B.4 Calculating A using a differential-equation system

We now combine Theorem 2 and Proposition 12 to compute A. We consider a productivity path $\boldsymbol{z}(t)$ connecting $\boldsymbol{z}^0$ and $\boldsymbol{z}$. The procedure repeatedly solves a linear system for changes in equilibrium variables and updates their levels. Throughout this subsection, prices, wages, and input shares refer to the compensated equilibrium.

We first present the differential-equation system and then the solution algorithm.

TFP-equivalent surplus

$$\frac{d\log A(t)}{dt} = \sum_{i\in C+N} \lambda_i^{\text{comp}}(t)\frac{d\log z_i(t)}{dt}, \qquad A(0)=1. \tag{56}$$

Input-output and prices

$$d\log \Omega_{ij} = (1-\theta_i)\begin{bmatrix} d\log p_j \mathbf{1}[j\in N] + d\log w_j \mathbf{1}[j\in R] \\ \quad -\sum_{k\in N}\Omega_{ik}\, d\log p_k - \sum_{r\in R}\Omega_{ir}\, d\log w_r \end{bmatrix}. \tag{57}$$

For entries of Ω that are always zero, including all entries in factor rows, we use the convention $d\log\Omega_{ij}=0$.

$$d\log \boldsymbol{p} = -\Psi^P\, d\log \boldsymbol{z} + \Psi^F\, d\log \boldsymbol{w}, \tag{58}$$

where $\boldsymbol{p}$ and $\boldsymbol{z}$ contain the produced commodities $i\in C+N$, Ψ^P is the corresponding produced-commodity block of the Leontief inverse $(I-\Omega)^{-1}$, and Ψ^F is its factor-content block.

Compensated supply and demand derivatives Let

$$L_r^{\text{comp}} = L_r^{\text{comp}}\left(\tfrac{\boldsymbol{w}}{A}, \boldsymbol{p}^c\right), \qquad \chi_r^{\text{comp}} = \chi_r^{\text{comp}}\left(\tfrac{\boldsymbol{w}}{A}, \boldsymbol{p}^c\right),$$

and define

$$\eta^L_{r,s} = \frac{\partial \log L_r^{\text{comp}}}{\partial \log(w_s/A)}, \quad \eta^{L,p}_{r,s} = \frac{\partial \log L_r^{\text{comp}}}{\partial \log p_s^c},$$
$$\eta^{\chi}_{r,s} = \frac{\partial \chi_r^{\text{comp}}}{\partial \log(w_s/A)}, \quad \eta^{\chi,p}_{r,s} = \frac{\partial \chi_r^{\text{comp}}}{\partial \log p_s^c}.$$

Then

$$d\log L_r^{\text{comp}} = \sum_{s\in R}\eta^L_{r,s}\left(d\log w_s - d\log A\right) + \sum_{s\in R}\eta^{L,p}_{r,s}\, d\log p_s^c, \tag{59}$$

$$d\chi_r^{\text{comp}} = \sum_{s \in R} \eta_{r,s}^{\chi} \left(d \log w_s - d \log A \right) + \sum_{s \in R} \eta_{r,s}^{\chi,p} \, d \log p_s^c. \tag{60}$$

Factor Domar weights

$$d \log \lambda_r^{\text{comp}} = d \log w_r + d \log L_r^{\text{comp}} - \sum_{r'} \lambda_{r'}^{\text{comp}} \left[d \log w_{r'} + d \log L_{r'}^{\text{comp}} \right]. \tag{61}$$

Using (59), this can be written in terms of $d \log w_s$, $d \log p_s^c$, and $d \log A$.

Commodity market clearing For every $i \in C + N + R$,

$$d \log \lambda_i^{\text{comp}} = \frac{1}{\lambda_i^{\text{comp}}} \left[\sum_{r \in R} d\chi_r^{\text{comp}} \mathbf{1}[i = c(r)] + \sum_{j \in C+N} \left(\Omega_{ji} \, d\lambda_j^{\text{comp}} + \lambda_j^{\text{comp}} \, d\Omega_{ji} \right) \right]. \tag{62}$$

Here $d\chi_r^{\text{comp}}$ comes from (60), $d\Omega$ from (57), and $d \log \lambda_j^{\text{comp}}$ from (61)–(62). Given $d \log A$ from (56) and $d \log z_i$, the equations (57)–(58)–(59)–(60)–(61)–(62) form a linear system that determines

$$d \log p_i, \quad d \log w_r, \quad d \log \Omega_{ij}, \quad d \log \lambda_i^{\text{comp}}$$

up to the choice of a numeraire.

Numerical integration algorithm

Step 0 (initialization, $t = 0$)

i. Solve the decentralized equilibrium at $t = 0$ to obtain

$$\boldsymbol{p}(0), \ \boldsymbol{w}(0), \ \boldsymbol{\Omega}(0), \ \boldsymbol{\lambda}(0).$$

ii. Set $A(0) = 1$.

Step 1 (shocks and grid)

i. Specify a path $\boldsymbol{z}(t)$ and compute $d \log z_i(t)/dt$.

ii. Choose a time grid $0 = t_0 < \cdots < t_K = t$ with $\Delta t_k = t_{k+1} - t_k$.

Step 2 (for each $k = 0, \ldots, K-1$) Given

$$\boldsymbol{p}(t_k),\ \boldsymbol{w}(t_k),\ \boldsymbol{\Omega}(t_k),\ \boldsymbol{\lambda}^{\text{comp}}(t_k),\ A(t_k),$$

proceed as follows.

2.1 Hulten step Compute

$$\left.\frac{d\log A}{dt}\right|_{t_k} = \sum_{i \in C+N} \lambda_i^{\text{comp}}(t_k)\ \left.\frac{d\log z_i}{dt}\right|_{t_k}.$$

2.2 Compensated demand and supply

i. Compute the relevant directional elasticities of compensated supply and demand:

$$\eta^{L}_{r,s},\ \eta^{L,p}_{r,s},\ \eta^{\chi}_{r,s},\ \eta^{\chi,p}_{r,s}$$

at that point (either analytically using Proposition 4 or by finite differences and simulation).

ii. Use (59) and (60) to express $d\log L_r^{\text{comp}}$ and $d\chi_r^{\text{comp}}$ as linear functions of $d\log w_s$, $d\log p_s^c$ and the known $d\log A\big|_{t_k}$.

2.3 General equilibrium

i. Use $p_r^c = p_{c(r)}$, and plug (59), (60), (57), and (58) into (61) and (62).

ii. This yields a linear system for

$$\frac{d\log p_i}{dt},\quad \frac{d\log w_r}{dt},\quad \frac{d\log \Omega_{ij}}{dt},\quad \left.\frac{d\log \lambda_i^{\text{comp}}}{dt}\right|_{t_k},$$

where $\left(d\log A/dt\right)\big|_{t_k}$ and $\left(d\log z_i/dt\right)\big|_{t_k}$ are known. Fix a numeraire and solve the system, omitting one redundant market-clearing equation.

2.4 Time update For each strictly positive component $x \in \{p_i, w_r, \Omega_{ij}, \lambda_i^{\text{comp}}, A\}$,

$$\log x(t_{k+1}) = \log x(t_k) + \left.\frac{d\log x}{dt}\right|_{t_k} \Delta t_k.$$

Structurally zero entries of Ω remain zero. Repeat until t_K.

C Appendix for Section 5

C.1 Proofs for Section 5

Proof of Proposition 5. All variables are evaluated in the compensated equilibrium, and we suppress the superscript comp. First we introduce producer-specific fixed factors X_f, whose effective supplies are X_f/A and whose shadow prices are w_f. These shadow prices may be negative. Without loss of generality, production is constant returns to scale in intermediate inputs, labor, and these fixed factors (since non-constant returns can always be modelled as constant returns with additional factors). We include payments to these auxiliary factors in household transfers so that the sum of household transfers must equal wedge revenues and payments to the fixed-factors:

$$\overline{T}(\boldsymbol{w}/A, \boldsymbol{p}^c) = \sum_i p_i y_i \left(1 - \frac{1}{\mu_i}\right) + \sum_f \frac{w_f X_f}{A} = \text{GDP} - \sum_r \frac{w_r L_r}{A}.$$

Here $\text{GDP} = \sum_{i \in C} p_i y_i$ and $\sum_i l_{ir} = L_r/A$. Differentiating the last equality and applying aggregate Shephard's lemma gives

$$\sum_{i \in C} p_i \, dy_i - \sum_{i,r} w_r \, dl_{ir} = \sum_r \frac{w_r L_r}{A} \, d\log A. \tag{63}$$

Differentiating producer i's technology and using its variable-input first-order conditions and fixed-factor shadow values gives

$$\frac{p_i}{\mu_i} dy_i = \frac{p_i y_i}{\mu_i} \, d\log z_i + \sum_j p_j \, dx_{ij} + \sum_r w_r \, dl_{ir} + \sum_f w_f \, d\left(\frac{X_{if}}{A}\right).$$

Sum across producers, cancel intermediate inputs by market clearing, and use $\sum_i X_{if} = X_f$ and $d(X_f/A) = -(X_f/A) d\log A$. Combining with (63) and dividing by GDP gives

$$\left(\sum_r \lambda_r + \sum_f \lambda_f\right) d\log A = \sum_i \frac{\lambda_i}{\mu_i} \, d\log z_i + \sum_i \lambda_i \left(1 - \frac{1}{\mu_i}\right) d\log y_i,$$

where $\lambda_r = (w_r L_r/A)/\text{GDP}$ and $\lambda_f = (w_f X_f/A)/\text{GDP}$ are factor-payment shares. The income identity above implies $\sum_r \lambda_r + \sum_f \lambda_f = 1 - \sum_i \lambda_i (1 - 1/\mu_i)$. Hence

$$\left[1 - \sum_i \lambda_i \left(1 - \frac{1}{\mu_i}\right)\right] d\log A = \sum_i \frac{\lambda_i}{\mu_i} \, d\log z_i + \sum_i \lambda_i \left(1 - \frac{1}{\mu_i}\right) d\log y_i. \tag{64}$$

Now represent each employment input (i,r) by an auxiliary sector with one unit of a producer-specific fixed factor, whose effective supply is $X_{(i,r)}/A = 1/A$. Its output is

$$y_{ir} = l_{ir}^{1+\alpha_{ir}}(1/A)^{-\alpha_{ir}} = \ell_{ir}^{\alpha_{ir}} l_{ir}, \qquad \ell_{ir} = A l_{ir}.$$

These auxiliary sectors generate no net household transfers, because wedge revenues exactly offset payments to the producer-specific fixed factors. Their prices therefore equal average labor cost rather than marginal cost. Their price, marginal cost, and implied wedge are

$$p_{ir} = \mathrm{ac}_{ir} = w_r A^{\frac{-\alpha_{ir}}{1+\alpha_{ir}}} y_{ir}^{\frac{-\alpha_{ir}}{1+\alpha_{ir}}} \qquad \mathrm{mc}_{ir} = w_r \frac{1}{1+\alpha_{ir}} (y_{ir}A)^{\frac{-\alpha_{ir}}{1+\alpha_{ir}}}, \qquad \mu_{ir} = \frac{p_{ir}}{\mathrm{mc}_{ir}} = 1+\alpha_{ir}.$$

Since the wedge offsets the payments to the producer-specific fixed factor, the sales share $p_{ir}y_{ir}/\mathrm{GDP}$ of these auxiliary producers is the original wage-bill share λ_{ir}. This representation therefore preserves GDP, household transfers, and the original producers' private input costs. Apply (64) to both original and auxiliary producers. The auxiliary sectors have no productivity shocks, and

$$d\log y_{ir} = (1+\alpha_{ir}) d\log \ell_{ir} - d\log A.$$

Their contribution to the right-hand side of (64) is

$$\sum_{i,r} \lambda_{ir} \left(1 - \frac{1}{\mu_{ir}}\right) d\log y_{ir} = \sum_{i,r} \lambda_{ir}\alpha_{ir}\, d\log \ell_{ir} - \sum_{i,r} \lambda_{ir} \frac{\alpha_{ir}}{1+\alpha_{ir}}\, d\log A.$$

The last term also appears on the left-hand side of (64) and cancels. We are left with

$$\left[1 - \sum_i \lambda_i \left(1 - \frac{1}{\mu_i}\right)\right] d\log A = \sum_i \frac{\lambda_i}{\mu_i}\, d\log z_i + \sum_i \lambda_i \left(1 - \frac{1}{\mu_i}\right) d\log y_i + \sum_{i,r} \lambda_{ir}\alpha_{ir}\, d\log \ell_{ir}.$$

□

Proof of Proposition 6. Start with an equilibrium with zero transfers. To prove this result, we construct lower and upper bounds with the same first-order change.

Let $C_r^0 = E_r^0/p_r^{c,0}$ denote initial consumption in location r. Hold households' consumptions and locations fixed, and let Z^{lb} adjust so that, holding labor supplies by location fixed, the production sector can deliver $\boldsymbol{C}^0$ using factor supplies $Z^{lb}\boldsymbol{L}^0$. This is always possible locally as long as every production function is smooth at the status quo.

Take this allocation, with producer prices and wages $(\boldsymbol{p}, \boldsymbol{w})$ that support this allocation given firms' first order conditions, and set $A^{lb} = 1/Z^{lb}$. This allocation is implementable using place-based compensations. To see this, set the common transfer T to zero and choose subsidies so that

$$1 - \tau_r = \frac{Z^{lb} w_r p_r^{c,0}}{w_r^0 p_r^c}.$$

Every household then faces its entire status-quo consumption menu:

$$\frac{Z^{lb} a_{hr} w_r}{(1-\tau_r) p_r^c} = \frac{a_{hr} w_r^0}{p_r^{c,0}}.$$

Thus, every household's original location and consumption are still optimal. Zero profits and market clearing imply $\sum_r p_r^c C_r^0 = Z^{lb} \sum_r w_r L_r^0$. Since $C_r^0 = (w_r^0 / p_r^{c,0}) L_r^0$, this identity also implies $\sum_r \tau_r p_r^c C_r^0 = 0$: the subsidies are budget balanced. Consequently, A^{lb} is a feasible solution for the problem defining A^{pb}. Hence, $A^{lb} \leq A^{pb}$.

For an upper bound, let A^{ub} be the largest $1/Z$ attainable by choosing households' consumptions and locations, together with all production plans, subject only to technological feasibility and

$$(c_h, l_h) \succeq_h (c_h^0, l_h^0) \qquad \text{for every } h.$$

Factor availability in location r is $ZL_r = Z \int a_{hr} \mathbf{1}[l_h = r]\, dh$. Every allocation implementable with place-based compensations is feasible for this unrestricted planner. Hence,

$$A^{lb} \leq A^{pb} \leq A^{ub}.$$

At the status quo, the first welfare theorem implies that the initial competitive allocation solves the unrestricted planner's problem. Thus, all three indices are equal to one initially.

Assume that the unrestricted planner's problem satisfies the regularity conditions for the constrained envelope theorem at the status quo. The initial competitive prices provide the planner's shadow values. The envelope theorem therefore gives

$$d \log A^{ub} = \frac{\sum_{i \in C+N} p_i^0 y_i^0\, d \log z_i}{\sum_r E_r^0} = \sum_{i \in C+N} \lambda_i^0\, d \log z_i,$$

where all differentials are evaluated at the status quo. Changes in household locations and production choices have no additional first-order effect on the optimized value. This also shows that, to a first-order, $d \log A^{lb} = d \log A^{ub}$, since the only difference between these allocations are changes that have no first-order effect. By the sandwich argument,

we have that $d \log A^{pb}$ also obeys Hulten's theorem.

$$\log A^{pb} \approx \log A \approx \sum_{i \in C+N} \lambda_i^0 \, \Delta \log z_i.$$

□

D Appendix for Section 6

D.1 Proofs for Section 6

Proof of Proposition 7. By Hulten's theorem (Hulten, 1978), to a first order we have:

$$\Delta \log Y \approx \sum_i \lambda_i(0) \Delta \log z_i + \sum_{r \in R} \lambda_r(0) \Delta \log L_r.$$

By Corollary 2,

$$\Delta \log A \approx \sum_i \lambda_i(0) \Delta \log z_i.$$

Subtracting the second equation from the first gives

$$\Delta \log Y - \Delta \log A \approx \sum_{r \in R} \lambda_r(0) \Delta \log L_r.$$

Using $\lambda_r(0) = w_r(0) L_r(0) / \sum_{r'} w_{r'}(0) L_{r'}(0)$, we have that to a first order,

$$\lambda_r(0) \Delta \log L_r \approx \frac{w_r(0) L_r(0)}{\sum_{r'} w_{r'}(0) L_{r'}(0)} \frac{\Delta L_r}{L_r(0)} = \frac{w_r(0)}{\sum_{r'} w_{r'}(0) L_{r'}(0)} \Delta L_r.$$

Summing across locations proves the proposition. □

Proof of Corollary 5. Here $Y(t)$ is the chained Divisia real-GDP index, normalized to one at the status quo, as discussed after Corollary 5. By Hulten's theorem (Hulten, 1978)

$$\log A^{\text{MFP}}(t) = \int_0^t \sum_i \lambda_i(s) \frac{d \log z_i}{ds} \, ds,$$

whereas Theorem 2 gives

$$\log A(t) = \int_0^t \sum_i \lambda_i^{\text{comp}}(s) \frac{d \log z_i}{ds} \, ds.$$

Since preferences rank allocations only by consumption and there is a common consumption good, household h ranks locations according to $(Za_{hr}w_r + T_h)/p^c$. Because Z and T_h are common across household h's location alternatives, this ranking is equivalent to ranking $a_{hr}w_r$. Consequently, location choices and labor supplies are independent of both Z and the compensating transfers. Hence,

$$l_h^{\text{comp}}(\boldsymbol{w}, \boldsymbol{p}^c) = l_h(\boldsymbol{w}, \boldsymbol{p}^c).$$

Furthermore, because there is only one common consumption good, Domar weights do not depend on the distribution of spending across locations. Hence, $\lambda_i^{\text{comp}}(s) = \lambda_i(s)$, for every i and s, which implies

$$A(t) = A^{\text{MFP}}(t).$$

We next show that $A^{MFP}(t) = Y(t)$. Hulten's theorem gives

$$\frac{d\log Y}{dt} = \sum_i \lambda_i(t)\frac{d\log z_i}{dt} + \sum_{r\in R}\lambda_r(t)\frac{d\log L_r}{dt}.$$

The second term is the reallocation effect. Consider a marginal mass dh of households moving from i to j. Since households choose locations to maximize nominal income, a marginal mover satisfies $a_{hi}w_i = a_{hj}w_j$. Its contribution to the reallocation effect is

$$\frac{\left(a_{hj}w_j - a_{hi}w_i\right)}{GDP}dh = 0.$$

Summing across all marginal flows implies $\sum_r \lambda_r(t)d\log L_r = 0$. Therefore,

$$\frac{d\log Y(t)}{dt} = \frac{d\log A^{\text{MFP}}(t)}{dt}.$$

Since both indexes equal one at $t = 0$, integrating gives $Y(t) = A^{\text{MFP}}(t)$, which proves the result. □

Proof of Proposition 8. For a continuum of households, interpret the minimum and maximum as the endpoints of the cross-sectional support of $A_h(t)$. Thus, for any b strictly between these endpoints, a positive measure of households has $A_h(t) < b$ and a positive measure has $A_h(t) > b$. For any $b > 0$, define

$$d_h(b) \equiv \max_r u_h\left(\frac{a_{hr}w_r(t)}{bp_r^c(t)}, r\right) - u_h^0.$$

By the definition of $A_h(t)$ and monotonicity in consumption, $d_h(b) > 0$ if $A_h(t) > b$, $d_h(b) < 0$ if $A_h(t) < b$, and $d_h(b) = 0$ if $A_h(t) = b$. By definition,

$$W\left(\frac{\boldsymbol{w}(t)}{b\boldsymbol{p}^c(t)}\right) - W\left(\frac{\boldsymbol{w}(0)}{\boldsymbol{p}^c(0)}\right) = \mathbb{E}[d_h(b)].$$

If $b < \min_h A_h(t)$, then $A_h(t) > b$ for every household, so $d_h(b) > 0$ for every household. Hence, $\mathbb{E}[d_h(b)] > 0$, and average utility at t, after dividing consumption by b, exceeds average utility in the status quo. Similarly, if $b > \max_h A_h(t)$, then $d_h(b) < 0$ for every household, so $\mathbb{E}[d_h(b)] < 0$. Therefore,

$$\min_h A_h(t) \leq A^W(t) \leq \max_h A_h(t).$$

Now fix any b strictly between these bounds and let

$$D_+ \equiv \mathbb{E}[d_h(b)\mathbf{1}\{A_h(t) > b\}], \qquad D_- \equiv -\mathbb{E}[d_h(b)\mathbf{1}\{A_h(t) < b\}],$$

which are both positive. Replace each utility representation u_h with

$$\widetilde{u}_h(c,r) = \begin{cases} D_- u_h(c,r), & A_h(t) > b, \\ D_+ u_h(c,r), & A_h(t) < b, \\ u_h(c,r), & A_h(t) = b, \end{cases} \qquad \text{for every } (c,r).$$

These positive rescalings leave ordinal preferences and $A_h(t)$ unchanged. Under the transformed cardinal utility functions, the change in average utility when consumption at t is divided by b is

$$\begin{aligned} &D_-\mathbb{E}[d_h(b)\mathbf{1}\{A_h(t) > b\}] + D_+\mathbb{E}[d_h(b)\mathbf{1}\{A_h(t) < b\}] \\ &\quad = D_-D_+ - D_+D_- = 0. \end{aligned}$$

Therefore, average utility after dividing consumption at t by b equals average utility in the status quo. By the definition of $A^W(t)$, this implies that $A^W(t) = b$. Because b was arbitrary, every value strictly between the bounds can be attained. □

D.2 Compensating variations in ratios and differences

In Section 6.2 we defined the sum of compensating variations as

$$S(t) \equiv -\int_h T_h(\boldsymbol{w}(t), \boldsymbol{p}^c(t))\, dh$$

The following lemma shows that $S(t)$ is equal to the maximum difference between total income and the sum of compensating incomes (minimum income to attain indifference with the status quo), evaluated at t equilibrium wages and prices.

Lemma 1 (Equivalent definition of $S(t)$)**.**

$$S(t) = \max_{\{(c_h, l_h)\}_{h\in H}} \left\{ \int_h \left[w_{l_h}(t)\, a_{hl_h} - p^c_{l_h}(t)\, c_h \right] dh : \; u_h(c_h, l_h) = u^0_h \; \forall h \right\}.$$

Proof. The minimizer in the definition of $T_h(\boldsymbol{w}, \boldsymbol{p}^c)$ must satisfy $u_h(c_h, l_h) = u^0_h$. Thus

$$T_h(\boldsymbol{w}, \boldsymbol{p}^c) = \min_{\substack{c_h, l_h, T_h \\ u_h(c_h, l_h) = u^0_h}} \left\{ T_h : \sum_r p^c_r c_h \mathbf{1}[l_h = r] \leq \sum_r a_{hr} w_r \mathbf{1}[l_h = r] + T_h \right\}.$$

For any (c_h, l_h), the budget constraint implies $T_h \geq p^c_{l_h} c_h - w_{l_h} a_{hl_h}$, so at the minimum $T_h = p^c_{l_h} c_h - w_{l_h} a_{hl_h}$. Hence

$$T_h(\boldsymbol{w}, \boldsymbol{p}^c) = \min_{\substack{c_h, l_h \\ u_h(c_h, l_h) = u^0_h}} \{ p^c_{l_h} c_h - w_{l_h} a_{hl_h} \},$$

and therefore

$$-T_h(\boldsymbol{w}, \boldsymbol{p}^c) = \max_{\substack{c_h, l_h \\ u_h(c_h, l_h) = u^0_h}} \{ w_{l_h} a_{hl_h} - p^c_{l_h} c_h \}.$$

Summing over h,

$$S(t) = \int_h \max_{\substack{c_h, l_h \\ u_h(c_h, l_h) = u^0_h}} \left\{ w_{l_h}(t) a_{hl_h} - p^c_{l_h}(t) c_h \right\} dh.$$

Because choices and payoffs are additively separable across households, maximizing the integral over all feasible profiles $\{(c_h, l_h)\}_{h\in H}$ (i.e. those that satisfy $u_h(c_h, l_h) = u^0_h \; \forall h$) is equivalent to choosing, for each h, a maximizer of the integrand, giving the stated result. □

Consider now an alternative definition of Kaldor-Hicks efficiency as the maximum

ratio — rather than the difference in terms of the numeraire — between aggregate income and the sum of compensating incomes, at t wages and prices:

$$A^{KH}(t) \equiv \max_{\{(c_h, l_h)\}_{h \in H}} \left\{ \frac{\int_h w_{l_h}(t)\, a_{hl_h}\, dh}{\int_h p^c_{l_h}(t)\, c_h\, dh} : \; u_h(c_h, l_h) = u^0_h \; \forall h \right\}.$$

Note that, in general, the allocations that maximize $S(t)$ do not coincide with those that maximize $A^{KH}(t)$. However, the following proposition shows that Kaldor-Hicks efficiency rises (i.e. $S(t) > 0$) if and only if $A^{KH}(t) > 1$. Therefore, $A^{KH}(t)$ inherits the same undesirable properties in general equilibrium as $S(t)$. For example, as discussed in Section 6.2, pure transfers cause Kaldor-Hicks efficiency to rise.

Proposition 15 (Equivalence in sign between $S(t)$ and $A^{KH}(t)$)**.**

$$\text{sign}\big(A^{KH}(t) - 1\big) = \text{sign}\big(S(t)\big).$$

Proof. We can write

$$\begin{aligned} \int_h \left[w_{l_h}(t)\, a_{hl_h} - p^c_{l_h}(t)\, c_h \right] dh &= \int_h w_{l_h}(t)\, a_{hl_h}\, dh - \int_h p^c_{l_h}(t)\, c_h\, dh \\ &= \left(\int_h p^c_{l_h}(t)\, c_h\, dh \right) \left(\frac{\int_h w_{l_h}(t)\, a_{hl_h}\, dh}{\int_h p^c_{l_h}(t)\, c_h\, dh} - 1 \right). \end{aligned}$$

Because prices and compensating consumption expenditures are positive,

$$\int_h p^c_{l_h}(t)\, c_h\, dh > 0$$

for every admissible profile. Thus, the sign of the left-hand side is the same as the sign of the term in parentheses. For every admissible profile, the sign of

$$\int_h \left[w_{l_h}(t)\, a_{hl_h} - p^c_{l_h}(t)\, c_h \right] dh$$

coincides with the sign of

$$\frac{\int_h w_{l_h}(t)\, a_{hl_h}\, dh}{\int_h p^c_{l_h}(t)\, c_h\, dh} - 1.$$

By Lemma 1, $S(t)$ is the maximum of the former over all feasible allocations, while $A^{KH}(t) - 1$ is the maximum of the latter over all feasible allocations. Since the sign equivalence holds allocation by allocation and both $S(t)$ and $A^{KH}(t) - 1$ are maxima over the

same feasible set, it follows that

$$S(t) > 0 \iff A^{KH}(t) > 1, \quad S(t) = 0 \iff A^{KH}(t) = 1, \quad S(t) < 0 \iff A^{KH}(t) < 1,$$

which is equivalent to $\text{sign}(A^{KH}(t) - 1) = \text{sign}(S(t))$. □

The virtue of defining A^{KH} as a ratio, rather than a difference, is that when real wages are pinned by technologies — i.e., they do not depend on labor supply $\boldsymbol{L}$ or on demand $\boldsymbol{\chi}$ — then $A^{KH}(t) = A(t)$. The assumption that real wages are independent of $\boldsymbol{L}$ and $\boldsymbol{\chi}$, that holds in the one-good economy of Example 1, is very strong. It implies that the real marginal product of labor is constant, which is to say that labor in each location can be converted into consumption goods in that location using a linear technology.

Proposition 16 (Equivalence between $A(t)$ and $A^{KH}(t)$)**.** *Suppose that real wages and relative prices in the decentralized equilibrium are independent of labor supply $\{L_r\}$ and demand $\{\chi_r\}$. Then $A^{KH}(t) = A(t)$.*

Proof. By Dinkelbach's theorem (Dinkelbach, 1967), the allocation that attains $A^{KH}(t)$ also maximizes

$$\int_h \left[\frac{w_{l_h}(t) a_{hl_h}}{A^{KH}(t)} - p^c_{l_h}(t) c_h \right] dh$$

subject to $u_h(c_h, l_h) = u^0_h$ for every h, and the maximized value is zero. By separability and Proposition 1, this allocation consists of the compensated choices evaluated at wages $\boldsymbol{w}(t)/A^{KH}(t)$. Therefore,

$$A^{KH}(t) = \frac{\sum_r w_r(t)\, L_r^{\text{comp}}\Big(\frac{\boldsymbol{w}(t)}{A^{KH}(t)}, \boldsymbol{p}^c(t)\Big)}{\sum_r E_r^{\text{comp}}\Big(\frac{\boldsymbol{w}(t)}{A^{KH}(t)}, \boldsymbol{p}^c(t)\Big)}. \tag{65}$$

This representation has the same form as Equation (11) in Theorem 1: there, $A(t)$ is characterized as the maximum proportional reduction in aggregate factor-augmenting productivity such that winners can compensate losers, with wages and prices evaluated in the compensated equilibrium. Here, $A^{KH}(t)$ is the maximum proportional reduction in wages in all locations such that it is possible for winners to compensate losers (based on prices and wages at t) and there is no money left over (sum of net transfers equals zero).

By assumption, equilibrium real wages and relative consumption prices do not depend on labor supply or demand. Hence the real wages and relative consumption prices in the compensated equilibrium coincide, up to a common nominal normalization, with

those in the decentralized equilibrium without transfers. Under this condition, Equation (11) reduces exactly to (65). Since (65) characterizes $A^{KH}(t)$ and Equation (11) characterizes $A(t)$, it follows that $A^{KH}(t) = A(t)$. □

D.3 Welfare effects of removing the Interstate Highway System

In this appendix, we repeat a counterfactual exercise in Allen and Arkolakis (2014), replacing the baseline bilateral trade-cost matrix with its no Interstate Highway System (IHS) counterpart. This removes the speed advantage associated with interstate highways while leaving the remaining road, rail, water, and air networks unchanged. We use the calibration described in the text without externalities, and consider three mobility elasticities: our baseline value of $\theta = 3$, and two sensitivity values, $\theta = 1/2$ and $\theta = 10$. For each θ, we recalibrate z_r and B_r so that the model exactly reproduces the same initial population and expenditure shares by county.

Table 1: Welfare effects of removing the IHS

	$\theta = 0.5$	$\theta = 3$	$\theta = 10$
Panel A. TFP-equivalent aggregate surplus, A	-1.59	-1.55	-1.47
Panel B. Consumption equivalent for average utility, A^W			
$u_h(c_r, r) = B_r c_r \varepsilon_{hr}$	-1.70	-1.66	-1.56
$u_h(c_r, r) = B_r c_r \varepsilon_{hr} / \max_{r'}\{\varepsilon_{hr'}\}$	-0.06	-1.42	-1.51
$u_h(c_r, r) = \exp(B_r c_r \varepsilon_{hr} / \max_{r'}\{\varepsilon_{hr'}\})$	-0.43	-0.64	-1.45
$u_h(c_r, r) = -\exp(-B_r c_r \varepsilon_{hr} / \max_{r'}\{\varepsilon_{hr'}\})$	-1.73	-1.68	-1.56
Minimum possible value	-13.20	-11.49	-8.94
Maximum possible value	4.47	2.81	1.09

Notes: All entries are 100 times log changes from the baseline to the no-IHS economy. All columns impose no externalities, $\alpha = 0$, and use $\sigma = 9$. The parameters $\{z_r\}$ and $\{B_r\}$ are set to target the same initial population and expenditure shares by county, together with $\sum_r (B_r c_r^0)^\theta / R = 1$. The bounds on A^W are from Proposition 8.

Panel A of Table 1 shows that removing the IHS reduces A by around 1.5–1.6%. The loss is larger when mobility elasticities are lower, consistent with Le Chatelier's principle. In contrast, Panel B shows that the consumption equivalent for average utility, A^W, depends on the cardinalization of utility. We consider the standard Fréchet cardinalization, normalization by each household's maximum taste draw, and convex and concave transformations of the normalized utility. These specifications represent the same ordinal preferences but, at the baseline $\theta = 3$, imply welfare losses ranging from 0.64% to 1.68%.

The results are more sensitive to the choice of cardinalization when mobility elasticities are lower. More generally, Proposition 8 shows that different cardinal representations of the same ordinal preferences can produce any consumption-equivalent change between the bounds reported in the last two rows. With $\theta = 0.5$, these range from a loss of roughly 13% to a gain of 4.5%, all consistent with the same underlying behavior.[46] As θ tends to infinity, households become homogeneous and their individual consumption equivalents coincide, so that the choice of cardinalization no longer matters.[47]

D.4 Average utility in dynamic discrete choice

Naively averaging utilities across households in a dynamic model is subject to the same criticisms as in a static model. However, if we restrict ourselves to comparing allocations for which all agents have the same preference ranking, then using average utilities is harmless since all households agree on the ranking, so a consumption-equivalent variation for average utilities is the same as one constructed agent-by-agent. We first describe this result in a general setting and then apply it to the Caliendo et al. (2019) framework.

Each household has a preference relation, $\succeq_h$, ranking state-contingent streams of consumption and location $\{c_t(s^t), l_t(s^t)\}_t$, where s^t is the history, including idiosyncratic tastes, up to date t. Suppose that, at each date $t-1$, conditional on moving to location r in the next period t before that date's taste shocks are observed, all households have the same preference rankings over the future starting at t. Then, the consumption-equivalent variation across two alternative consumption-location streams that both start at location r in t is the same for every agent, regardless of that agent's identity and where that agent is in period $t-1$. Specifically, letting u_h denote a utility function representing $\succeq_h$, the consumption-equivalent variation A_r that solves the following indifference condition

$$u_h\left(\bar{c}_{t-1}(s^{t-1}), \bar{l}_{t-1}(s^{t-1}), A_r c_t, r, A_r c_{t+1}(s^{t+1}), l_{t+1}(s^{t+1}), \dots\right) = u_h\left(\bar{c}_{t-1}(s^{t-1}), \bar{l}_{t-1}(s^{t-1}), c'_t, r, c'_{t+1}(s^{t+1}), l'_{t+1}(s^{t+1}), \dots\right), \quad (66)$$

[46] An odd feature of the common $u_h(c_r, r) = B_r c_r \varepsilon_{hr}$ cardinalization is that average utility diverges when the mobility elasticity θ is less than or equal to one. This is because the Fréchet taste parameters ε_{hr} are sufficiently fat-tailed that their expectation is infinite. However, we can still calculate A^W using (33) if we cancel the diverging constants from both sides before taking a limit where θ falls below one. So, while the level of average utility is infinite when $\theta < 1$, there is still a way to compute the consumption-equivalent change. The alternative cardinalization, where you normalize by the maximum value of ε_{hr}, does not have this divergence issue.

[47] As $\theta \to \infty$, we recover the Allen–Arkolakis homogeneous-agent case, with a welfare loss of 1.1%. In this case, $A^W = A$; see Appendix A.2 where we discuss aggregate welfare in economies with homogeneous agents.

is the same for every agent. This calculation produces a consumption-equivalent variation A_r for each location r. Since A_r is the same number for every household, it is also the consumption-equivalent variation if we use a utilitarian or average utility welfare function (since all households have the same preference ranking over these restricted allocations).

In Caliendo et al. (2019), each agent h's preferences at time t are described by:

$$v_{ht}^r = \log c_t^r + \max_s \left\{ \beta V_{t+1}^s - \tau^{rs} + \epsilon_{ht}^s \right\}, \qquad V_t^r = \mathbb{E}_{\epsilon_{ht}}[v_{ht}^r], \tag{67}$$

where r and s index locations and τ^{rs} is a moving cost. The expectation defining V_t^r is over the vector of taste shocks $\epsilon_{ht} = \{\epsilon_{ht}^s\}_s$ for each location s. With independent, mean-zero, type-I extreme-value draws of ϵ's across destinations and time, any household choosing at $t-1$ to enter market r at t has the same expected continuation utility from t onward, evaluated before observing its date-t tastes:

$$V_t^r = \log c_t^r + \nu \log \left(\sum_s \exp \left\{ \beta V_{t+1}^s - \tau^{rs} \right\}^{\frac{1}{\nu}} \right), \tag{68}$$

where $\nu > 0$ is a scale parameter.

Applying (66), it is straightforward to show that the consumption-equivalent A_r solves

$$\begin{aligned} &\log \bar{c}_{t-1}^{\bar{l}_{h,t-1}} - \tau^{l_{h,t-1}r} + \epsilon_{h,t-1}^r + \beta \left[V_t^r + \frac{\log A_r}{1-\beta} \right] \\ &\quad = \log \bar{c}_{t-1}^{\bar{l}_{h,t-1}} - \tau^{l_{h,t-1}r} + \epsilon_{h,t-1}^r + \beta V_t^{r\prime}, \end{aligned} \tag{69}$$

The household-h specific terms cancel out, leaving:

$$\log A_r = (1-\beta)(V_t^{r\prime} - V_t^r).$$

This is the welfare change, for each r, reported by Caliendo et al. (2019) (right before equation 31 in Appendix A.2 of their paper). It does not depend on the households' identity, h, or the date-$t-1$ location $l_{h,t-1}$, and is invariant to monotone transformations of utility.

The restriction on the comparisons in (66) is important. Fixing consumption and location at $t-1$ alone does not fix the market entered at t. Suppose instead that the baseline continuation starts in r at t and the counterfactual continuation starts in s. Holding consumption and location at $t-1$ fixed, the household prefers the counterfactual if and only

if

$$\beta\left(V_t^{s\prime} - V_t^r\right) > \tau^{l_{h,t-1}s} - \tau^{l_{h,t-1}r} + \left(\epsilon_{h,t-1}^r - \epsilon_{h,t-1}^s\right). \tag{70}$$

For $s = r$, the right-hand side is zero. But if the two locations are different, $s \neq r$, then the comparison depends on the household's date-$t-1$ tastes and possibly its date-$t-1$ location, so households can disagree. Moreover, the consumption equivalent from a fixed starting market r need not equal that from another starting market: $A_r \neq A_s$.

As discussed in the Remark on page 12, we can also apply our notion of welfare to measure surplus for any subset of agents. If one applies our measure to the same comparison as (66) for an infinitesimal agent, then one obtains the same consumption-equivalent welfare as A_r in (69).

However, our framework allows us to measure aggregate surplus without requiring this type of homogeneity of agents. In particular, we can measure collective surplus for all agents across all regions and allowing for changes in the initial period in the counterfactual. For a group of heterogeneous households, our measure uses a common endowment contraction and balanced-budget transfers to keep each household indifferent to its own status quo, allowing compensated equilibrium prices to adjust. It does not require households to agree on rankings or consumption equivalents, and can accommodate changes in choices from date t onward and comparisons across markets.