NBER WORKING PAPER SERIES

AGGREGATE PRODUCTIVITY WITH HETEROGENEOUS AGENTS

David Baqaee
Ariel Burstein

Working Paper 34176
http://www.nber.org/papers/w34176

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2025, Revised January 2026

We thank Rodrigo Adao, Manuel Amador, Andrew Atkeson, Natalie Bau, Pablo Fajgelbaum, Joao Guerreiro, Valentin Haddad, Keith Head, Oleg Itskhoki, Dirk Krueger, Yuhei Miyauchi, Elisa Rubbo, Aleh Tsyvinski, Gianluca Violante, Mike Waugh, and Pierre-Olivier Weill for their comments. This paper subsumes Aggregate Welfare and Output with Heterogeneous Agents, first circulated in 2022. We thank Lautaro Adamovsky for exceptional research assistance. We acknowledge financial support from NSF grant No. 2343691. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2025 by David Baqaee and Ariel Burstein. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Aggregate Productivity with Heterogeneous Agents
David Baqaee and Ariel Burstein
NBER Working Paper No. 34176
August 2025, Revised January 2026
JEL No. A10, C0, D0, D3, D50, D60, E0, F0, H0

ABSTRACT

We develop a welfare-based measure of aggregate productivity for economies with heterogeneous households. For any change in the economic environment, we define the associated change in aggregate productivity as the largest shift in total factor-augmenting productivity that makes it feasible to leave every household at least as well off as under the status quo allocation. This construction maps arbitrary shocks to the economy into a TFP-equivalent change. In the absence of household heterogeneity, it nests the standard notions of aggregate TFP and consumption-equivalent welfare changes. It can also be interpreted as a general-equilibrium analogue of cost-benefit analysis, increasing whenever the gains to winners are more than sufficient to compensate the losers. We show how to port results that hold for aggregate TFP in representative agent settings, like Hulten's theorem and Harberger triangles, to this setting. We characterize changes in aggregate productivity in terms of observables, including expenditures and price elasticities, and apply our measure to study the effects of productivity shocks, the costs of misallocation, and the impact of trade shocks, both with and without costly redistribution.

David Baqaee
University of California, Los Angeles
Department of Economics
and CEPR
and also NBER
baqaee@econ.ucla.edu

Ariel Burstein
University of California, Los Angeles
Department of Economics
and NBER
arielb@econ.ucla.edu

1 Introduction

Two central themes in economics are efficiency — how big is the economic pie — and equity — how it ought to be split. The normative question of how resources should be divided among people is controversial and subjective. For this reason, economists try to study efficiency in isolation from equity when possible. With a single agent, efficiency is trivially separated from equity because distributional issues never arise. With heterogeneous consumers, however — especially when consumers differ in tastes or face different prices — the two questions are harder to separate.

In academic and theoretically-oriented settings, outcomes across heterogeneous agents are often summarized using social welfare functions. As is well-known, this approach, which was pioneered by Bergson (1938) and Samuelson (1947), mixes efficiency and equity considerations. Social welfare functions automatically take a stance not just on the size of the economic pie, but also on its optimal division. This means that social welfare functions depend on subjective Pareto-weights and are not invariant to monotone transformations of utility functions. This makes them unattractive for use in practical policy design, especially if heterogeneity in tastes is important (i.e. in cases where we cannot use the same cardinal utility function to represent every agent's preferences).[1]

In more policy-oriented circles, changes in aggregate efficiency are typically measured by summing up willingness-to-pay and cost-benefit analysis: "can the winners (potentially) compensate the losers and still be better off?" This may be done explicitly, by estimating and summing compensating variations, or implicitly, by using market-based quantity indices like real GDP (which, under appropriate assumptions, are equivalent to summing compensating variations). These metrics measure the change in aggregate efficiency by the amount of money (at constant prices) left-over after winners compensate losers. Such measures are popular because they are objective — taking the view that a dollar is a dollar, regardless of who earns it — and easy to communicate (Harberger, 1971).

However, it is well-known that standard cost-benefit type measures like real GDP and the sum of compensating variations have theoretical problems. A major issue is that it may not be feasible for winners to compensate the losers even though the sum of compensating variations is positive. This is either because lump-sum transfers are not available (e.g. Antràs et al., 2012 and Schulz et al., 2023) or because post-shock prices change if winners try to compensate losers (e.g. Boadway, 1974).

[1] A common way to neutralize distributional motives in social welfare functions is to set Pareto weights equal to Negishi weights (those that rationalize the decentralized status quo allocation). This removes redistributive motives only to a first-order, it does not eliminate them nonlinearly.

In this paper, we extend the cost-benefit approach to deal with these shortcomings, building on ideas from Allais (1979), Debreu (1951), and Luenberger (1996). Our measure also generalizes the popular Lucas (1987) consumption-equivalent beyond the one-agent setting. For any change in the economic environment, we define the associated change in aggregate productivity as the largest shift in total factor-augmenting productivity that makes it feasible to leave every household at least as well off as under the status quo allocation. This measures the resources left over after winners compensate the losers: aggregate productivity rises whenever the same welfare as in the status quo can be achieved with less resources. We do not refer to this as a measure of aggregate welfare to distinguish it from a social welfare function that embeds normative judgements about interpersonal utility comparisons. In our baseline, we assume that compensations can be done using lump-sum transfers, though we extend this to allow for costly and distortionary compensations later.

Our measure answers a purely positive question. Unlike social welfare functions, it does not introduce free parameters such as Pareto weights and is invariant to monotone transformations of utility (i.e. it depends only on ordinal properties of preference relations). Moreover, the measure does not take a normative stance on the optimal distribution of resources across individuals beyond compensating every agent relative to the status quo. For example, if aggregate efficiency increases and there are extra resources left over after everyone has been compensated, our measure takes no stance on who should get those resources.

The numerical value of our efficiency measure is interpretable and, in the special case of a single agent, collapses to familiar objects like real GDP, the sum of compensating variations, and the consumption-equivalent variation measure of Lucas. However, it avoids the limitations and paradoxes of these measures. We characterize aggregate productivity in terms of observables like expenditures and price elasticities, and we use it to generalize well-known representative-agent results to settings with heterogeneous-agents. This paper has two stand-alone companions —Baqaee and Burstein (2025a) and Baqaee and Burstein (2025b) — where we apply this framework to answer two different questions: the cost of misallocation due to financial market incompleteness (in both closed and open-economies) and changes in aggregate productivity in random utility models with discrete choice. In both applications, household heterogeneity is a central feature of the problem.

The structure of the paper is as follows. In Section 2, we set up the environment and define our measure of aggregate productivity (with lump-sum transfers). For comparison, we also review some popular alternative measures of efficiency: real GDP (using a Divisia or chain-weighted index), Kaldor-Hicks efficiency (using the sum of compensating variations), and the welfare of a positive representative agent (if such an agent exists).

In Section 3, we show that, under some assumptions, aggregate productivity can be calculated by solving the equilibrium of a fictional representative agent economy. The utility function of the fictional representative agent equals the change in aggregate productivity. This result makes it straightforward to port tools and methods from representative-agent economies to analyze aggregate productivity with heterogeneous agents. This forms the basis for many of the other results in the paper. In this section, we establish an important benchmark result: if all households have identical homothetic preferences and face the same relative prices, our measure of aggregate productivity (with lump-sum transfers) coincides with those alternative measures. Outside of these common but restrictive assumptions (which rule out preference heterogeneity and incomplete markets), however, the measures generally differ.

In Section 4 we restrict attention to perfectly competitive economies without distortions. We show that in such settings and to a first-order approximation, Hulten's theorem applies to our measure of aggregate productivity unaltered. That is, to a first-order and in competitive models, our measure of aggregate productivity coincides with the Solow (1957) residual and Kaldor-Hicks efficiency. However, this equivalence breaks down beyond a first-order approximation. Even in competitive economies, for large shocks, we show that real GDP, the Solow residual, and the sum of compensating variations have pathological properties that our measure does not have. We derive a nonlinear version of Hulten (1978) that applies to our measure and use this to extend the nonlinear characterizations in Baqaee and Farhi (2019c) to economies with heterogeneous agents. We show that changes in aggregate productivity depend only on expenditure shares and price elasticities. We also generalize the sufficient-statistics of Arkolakis et al. (2012), developed for single-agent economies, to quantify the gains from trade relative to autarky in economies with heterogeneous agents.

In Section 5 we consider distorted economies, and derive versions of Hsieh and Klenow (2009), Petrin and Levinsohn (2012), Harberger (1954, 1964), and Baqaee and Farhi (2020) that apply to economies with heterogeneous agents. We show that, once prices are not equal to marginal costs, then real GDP and Kaldor-Hicks efficiency are not reliable measures of aggregate efficiency even to a first-order. In particular, they can rise or fall in response to pure redistributions even as we move from one Pareto-efficient allocation to another. We derive a version of the famous Harberger triangles formula that can be used to quantify misallocation with heterogeneous agents, and show that there is a sense in which misallocation losses in the heterogeneous agent model are lower than in representative agent models. This is because our Harberger triangles formula discards dispersion in wedges due to difference in average wedges paid by each household. Differences in the average wedge by household are equivalent to lump-sum transfers and do not imply (Pareto) inefficiency.

In Section 6 we consider economies with costly redistribution (i.e. without lump-sum transfers). We show that, starting in perfect competition, the change in aggregate productivity is, to a first-order, the same as Hulten (1978) (under some mild assumptions). To a second-order, the change in aggregate productivity is equal to what would have happened with lump-sum transfers (characterized in Section 4) minus the additional Harberger triangles (characterized in Section 5) caused by inefficient redistribution (which are zero if lump-sum transfers are available). We provide a worked-out example showing how limited redistributive tools raise the costs of moving to autarky compared to when lump-sum transfers are available. We end the section with a quantitative example studying how the rise of China affected aggregate productivity in the United States. We show that the increase in productivity for the U.S. due to the rise of China depends on the ease with which workers can move across sectors, and the range of redistributive tools available. Whereas the change is positive when workers can move across sectors or if lump-sum transfers are available, it is negative if workers are restricted to working within narrow industries and redistribution is impossible or costly (highly distorting).

Related literature. Our approach to measuring aggregate productivity is related to willingness-to-pay based measures, which have a very long history in economics. (For example Dupuit, 1844; Hicks, 1939; Kaldor, 1939). Our approach is also related to the notion of social surplus in Allais (1979), the coefficient of resource utilization in Debreu (1951, 1954), the measure of efficiency in Farrell (1957), and the benefit function in Luenberger (1996). Our contribution relative to these works is to provide a characterization without assuming Pareto efficiency or lump-sum transfers and applying these types of measures to modern models.

Our paper is also related to cost-benefit analysis, typically performed by using the sum of compensating variations, as in Harberger (1971), and related ideas like the marginal value of public funds (Hendren and Sprung-Keyser, 2020). The idea behind these measures is to ask: "after the winners compensate the losers using lump-sum transfers, is there still money left on the table?" Our measure of productivity coincides with these measures when the equilibrium is perfectly competitive, the consumption-possibility set is linear, and lump-sum transfers are available. However, outside of these cases, our measures is different. First, if prices are not equal to marginal costs, then pure transfers can raise or lower real GDP and the sum of compensating variations, even if allocations are

Pareto-efficient. In contrast, our measure of productivity does not increase unless a Pareto improvement is possible. Second, if the consumption-possibility set is nonlinear, then as shown by Boadway (1974), a pure transfer between agents can cause the sum of compensating variations to exceed zero. Our measure does not have these property. It only rises if there is a potential Pareto improvement. Third, unlike the sum of compensating variations, our measure need not presuppose that lump-sum transfers are available. In this sense, our approach shares strong similarities to Schulz et al. (2023), who generalize the sum of compensating variations to allow for limited redistribution.[2]

Our paper complements and differs from Schulz et al. (2023) in many ways, the most important being a difference in focus. They consider economies with a single consumption good, focusing their attention on a mechanism design problem where lump-sum taxes are unavailable because of asymmetric information. Although our formalism and definitions can be applied to such economies, we do not focus on these issues. Instead, we focus on allowing for multiple goods and heterogeneity in preferences and relative prices faced by consumers. This means that even with perfect information and lump-sum transfers, there are interesting questions about how to aggregate across consumers that consume and value different goods.[3]

As mentioned above, a different approach to aggregation is to use a social welfare function to evaluate outcomes. A prominent example is the behind the veil-of-ignorance measure of Harsanyi (1955). Social welfare functions are by far the most common approach in the modern literature to aggregating across heterogeneous agents.[4] Our paper, which instead looks for and quantifies the potential for Pareto improvements (i.e. compensating everyone and looking to see if resources are left over), studies an alternative question that the one analyzed by this methodology.

[2]In response to a shock, they consider a tax reform that makes households indifferent to the status quo and then measure the monetary value of aggregate welfare gains or losses by the fiscal surplus from this reform.

[3]A related approach, Auerbach and Kotlikoff (1987), quantifies aggregate efficiency in overlapping-generations economies by using lump-sum transfers to keep generations before a specified date at their status quo utility level and increase the utility of all cohorts after that date by a common amount. Our measure is different because it keeps every agent indifferent to the status quo, allows for limited redistribution, and measures efficiency in terms of resource-savings instead. It would be interesting to apply our measure to an overlapping-generations economy.

[4]There is a branch of the literature that assumes observed allocations can be rationalized by maximizing some social welfare function within some parametric class, estimates this function, and uses it to conduct policy analysis (see Heathcote and Tsujiyama, 2021 and the references therein). This is equivalent to assuming there exists a normative representative agent: a hypothetical single decision-maker whose utility function is maximized by observed allocations (Chapter 4 Mas-Colell et al., 1995). Our approach is different since we do not need to assume the existence of either a positive nor normative representative agent. Furthermore, even if a normative representative agent exists, there is nothing to say that its preferences should be privileged over any other social welfare function (see Example 7 below).

Following in the social-welfare-function tradition, a recent set of papers, including Bhandari et al. (2021), Dávila and Schaab (2022, 2023), and Donald et al. (2023) provide approximate decompositions of changes in social welfare functions. Our goal in this paper is different: we do not provide decompositions of social welfare functions, but instead, define and characterize aggregate productivity directly as an answer to a purely positive question. The decompositions in the papers mentioned above contain components the authors refer to as capturing efficiency. However, since our objective is different, our notion of productivity is also generically different to the efficiency components in these papers (even to a first-order). Defining productivity directly, instead of as part of an infinitesimal decomposition, is useful because it means that we can also study large changes.[5]

In terms of the tools and methods, our paper is most closely related to the literature that studies the macroeconomic consequences of microeconomic productivity changes and wedges. For productivity changes, this includes Gabaix (2011), Acemoglu et al. (2012), Baqaee and Farhi (2019c) and others. For wedges, this includes Harberger (1954), and more recently, Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Bigio and La'O (2016), Liu (2017), Baqaee and Farhi (2020), among others. We relax the assumption typically maintained in both of these literatures that households have common preferences and face common prices.[6]

Finally, because we use the gains from trade as one of our examples, our paper is also related to the gains from trade with heterogeneous agents. Much of the work on international trade with heterogeneous agents focuses on the distributional effects of trade. Some examples of papers that also calculate aggregate welfare are Antras et al. (2017) and Galle et al. (2023) (using an Atkinson (1970)-style social welfare function with inequality aversion), Kim and Vogel (2020) (using the sum of compensating variations), and Rodríguez-Clare et al. (2022) (using a population-weighted average of welfare gains across regions), all of which differ from our measure of aggregate efficiency for reasons already discussed.

[5] Whereas infinitesimal changes in our measure of efficiency can be integrated to study large changes, integrals of components in a decomposition of social welfare are path-dependent. To see this point, suppose we approximately decompose changes in some function $y = f(x_1, x_2)$ into $dy \approx (\partial f/\partial x_1)dx_1 + (\partial f/\partial x_2)dx_2$. Then we can write non-infinitesimal changes as $\Delta y = \int(\partial f/\partial x_1)\Delta x_1 + \int(\partial f/\partial x_2)\Delta x_2$ but, unless $f(x_1, x_2)$ is linear in x_1 and x_2, the size of each component of this nonlinear decomposition depends on the arbitrary path of integration.

[6] One exception is Bornstein and Peter (2024), who study misallocation with differences in tastes and markups across households. In their setting, symmetry and the law-of-large numbers implies that every households' problem is identical despite the fact that households have different preferences.

2 Setup

In Section 2.1, we set up a flexible general-equilibrium framework, and in Section 2.2, we define aggregate productivity in this environment. In Section 2.3, we define some other measures of aggregate activity/welfare (real GDP, the sum of compensating variations, and consumption-equivalents for a positive representative agent) that are popular in the literature so that we can compare them.

2.1 Economic Environment

We consider Walrasian equilibrium with heterogeneous agents, arbitrary neoclassical production functions, and arbitrary distorting wedges. We describe the households' problem, followed by the producers', and then the resource constraints.

Households. Households are indexed by $h \in \{1, ..., H\}$. Agent h has ordinal preferences $\succeq_h$ over commodity vectors $\boldsymbol{c}_h \in \mathbb{R}^N$, where N is the number of goods. Assume preferences are represented by utility functions $u_h(\boldsymbol{c}_h)$.[7] A *consumption allocation* is a matrix $\boldsymbol{c} \in \mathbb{R}^{H\times N}$ whose hth row, denoted by $\boldsymbol{c}_h$, equals the consumption vector of agent h. Each household maximizes utility subject to a budget constraint

$$\max\ u_h(\boldsymbol{c}_h) \text{ such that } \sum_i p_i c_{hi} \leq \sum_f \omega_{hf} w_f L_f + T_h, \tag{1}$$

where the left-hand side is total expenditures and the right-hand side is total income. As in Arrow-Debreu, commodities could be indexed by time and state of nature. On the left-hand side, p_i is the price of i and c_{hi} is the quantity of good i purchased by household h. On the right-hand side, households derive income from factors and lump-sum transfers. Households h owns a share ω_{hf} of factor f, where w_f is the wage and L_f is the total quantity of factor f. Lump-sum transfers are T_h.

Producers. Producer i chooses its inputs to minimizes costs

$$\min \sum_j p_j y_{ij} + \sum_f w_f l_{if}, \text{ such that } y_i = z_i F_i\left(\{y_{ij}\}, Z\{l_{if}\}\right), \tag{2}$$

where y_i is the quantity of output, F_i is a constant-returns production function, y_{ij} are intermediate inputs used by i produced by j, and l_{if} are primary factors used by i. The

[7] That is, for each household $u_h(\boldsymbol{c}_h) \geq u_h(\boldsymbol{c}'_h)$ if, and only if, $\boldsymbol{c}_h \succeq_h \boldsymbol{c}'_h$. We assume that preferences are continuous, convex, and locally nonsatiated.

scalar Z is an aggregate factor-augmenting productivity shifter. The assumption that F_i has constant-returns is without loss of generality, since we can capture decreasing returns using producer-specific factors. The parameter z_i is a Hicks neutral productivity shifter. The price of i is equal to an exogenous markup or tax, $\mu_i > 0$, times i's marginal cost of production

$$p_i = \mu_i mc_i. \tag{3}$$

That is, the price of i is inclusive of the wedge on i's output.

Remark (Buyer-seller-specific productivity and wedges). Although we assume that z_i is Hicks neutral and wedges are on gross output only, both of these assumptions are made without loss of generality. This is because we can recreate buyer-seller productivity changes and wedges by relabeling. Specifically, we can treat firm or household i's purchases of an input from j as a distinct good (made linearly using j's output). A productivity shock or a wedge on this good is then isomorphic to a buyer-seller specific productivity shock or wedge. We make the assumption that z_i is Hicks neutral and assume all wedges take the form of taxes on gross output to simplify the notation.

Remark (Non-Walrasian Economies and Endogeneous wedges). Although we define a Walrasian equilibrium, the presence of the wedges allow us to replicate non-Walrasian economies. In particular, the wedges μ_i could themselves be functions of other endogeneous variables. This allow us to capture, for example, nominal rigidities (e.g. Rubbo, 2020), variable markups (e.g. Baqaee et al., 2024), and financial market incompleteness (e.g. Baqaee and Burstein, 2025b).

Resource constraints. The resource constraint for goods and factors is

$$\sum_j y_{ji} + \sum_h c_{hi} \leq y_i, \qquad \text{and} \qquad \sum_i l_{if} \leq z_f L_f, \tag{4}$$

where z_f, when f indexes a factor, controls the endowment of efficiency units of factor f. Finally, net transfers across households are equal to the revenues generated by the wedges:

$$\sum_h T_h = \sum_i p_i y_i \left(1 - \frac{1}{\mu_i}\right). \tag{5}$$

We now define a general equilibrium with wedges.[8]

Definition 1 (Decentralized Equilibrium with Wedges)**.** A *decentralized equilibrium with*

[8]This notion of general equilibrium is the same one used by Baqaee and Farhi (2020), extended to allow for multiple households.

wedges is the collection of prices and quantities such that: (1) the price of each good i equals its marginal cost times a wedge μ_i; (2) each producer chooses quantities to minimize costs taking prices as given; (3) each household chooses consumption quantities to maximize utility taking prices, consumption taxes, and income as given; (4) net transfers across households are equal to wedge revenues; (5) all resource constraints are satisfied.

2.2 Definition of Aggregate Productivity

We now define our measure of aggregate productivity. Index exogenous parameters of the economy (productivities, wedges, factor ownership, and transfers) by a scalar t and let $t = 0$ denote the status quo allocation. For any equilibrium price or quantity X, we write $X(t)$ to denote its dependence on the exogenous parameters.[9] Without loss of generality, we assume that Z is constant and equal to one as a function of t.[10]

We take the status quo, $t = 0$, to be the observed equilibrium allocation (i.e. parameter values under which the model is mapped to the data). The status quo consumption allocation is $\boldsymbol{c}(0) \in \mathbb{R}^{H\times N}$ — note that in a dynamic or stochastic model, $\boldsymbol{c}(0)$ is the entire stochastic process for consumption given initial parameter values, not the consumption realizations in the first period of the model.

Let $\mathcal{C}(t, Z)$ denote the set of consumption allocations that can be supported as part of an equilibrium given technologies $\boldsymbol{z}(t)$, wedges $\boldsymbol{\mu}(t)$, some lump-sum transfers, and factor-augmenting technology level Z:

$$\mathcal{C}(t, Z) = \left\{ \boldsymbol{c} \in \mathbb{R}^{H\times N} : \text{ there exist transfers supporting } \boldsymbol{c} \text{ as equilibrium, satisfying (1)-(5)} \right\}.$$

This is the set of consumption allocations that can be attained via lump-sum transfers and depends on t because technologies $\boldsymbol{z}(t)$ and wedges $\boldsymbol{\mu}(t)$ are indexed by t. It depends on factor-augmenting productivity shifter Z because the set $\mathcal{C}(t, Z)$ changes shape and shifts out as Z rises. In the absence of wedges, $\boldsymbol{\mu}(t) \equiv \mathbf{1}$, the second welfare theorem implies that the consumption possibility set, defined above, is the set of Pareto efficient consumption allocations.

Definition 2 (Aggregate Productivity)**.** *Aggregate productivity* at t, given lump-sum transfers, is the maximum contraction of factor-augmenting productivity such that every agent

[9] In the case of multiple equilibria, we assume there is an equilibrium selection mechanism. The nature of this equilibrium selection mechanism is not relevant for aggregate productivity, because aggregate productivity is unique given t and the status quo.

[10] Changes in aggregate factor-augmenting productivity as a function of t can be captured by uniform changes in the efficiency units of factors, z_f.

can be kept at least indifferent to the status quo allocation. Formally,

$$A(t) \equiv \max \left\{ Z \in \mathbb{R} : \text{there is } \boldsymbol{c} \in \mathcal{C}(t, 1/Z) \text{ and } \boldsymbol{c}_h \succeq_h \boldsymbol{c}_h(0) \text{ for every } h \right\}. \tag{6}$$

If $A(t) > 1$, then the economy's consumption possibility set, given lump-sum transfers, at t — $\mathcal{C}(t, 1)$ — contains a strict Pareto-improvement relative to the status quo, and if $A(t) < 1$, then at every point in $\mathcal{C}(t, 1)$, at least one agent is worse off than in the status quo. The cardinal value of $A(t)$ is interpretable: it converts the shock at t into an equivalent change in total factor productivity that brings everyone back to the status quo. For concreteness, say, $A(t) = 1.01$, then this means that it is possible to make everyone at least as well off as in the status quo and discard around 1% of every factor (or more precisely, $1 - 1/A$ percent). Agents may not be consuming the same bundle as in the status quo after they are compensated — we only require that they be indifferent to the status quo. If $\boldsymbol{\mu}(t) = \mathbf{1}$, and $\boldsymbol{z}(t) = \boldsymbol{z}(0)$, then $A(t)$ is the Debreu (1951) *coefficient of resource utilization*, measuring the distance of the status-quo from the Pareto efficient frontier, where there are no distortions, in units of factor endowments.[11]

The measure in Definition 2 has some desirable properties: (1) it answers a counterfactual question about observable phenomena with interpretable units. That is, $A(t)$ is invariant to monotone transformations of utility functions, and only relies on ordinal properties of preference relations. (2) This measure does not take a stance on how social surplus or losses should be divided among agents. That is, while we can evaluate aggregate productivity for some counterfactual technologies and wedges, we do not attempt to pick a specific feasible allocation as being socially "optimal." (3) This definition can easily be modified to environments with imperfect redistributive tools. Indeed, Section 6 is devoted to this topic. (4) As we show below, this measure coincides with traditional measures like real output (GDP), Kaldor-Hicks efficiency, and Lucas consumption-equivalent variation in cases where those measures are well-behaved. In this sense, $A(t)$ provides a way to extend these measures to capture cases with preference heterogeneity.[12]

[11] Proposition 11 in the appendix shows that, if $\boldsymbol{\mu}(t) = \mathbf{1}$, defining $A(t)$ using the set of equilibrium consumption allocations $\mathcal{C}(t, Z)$ is equivalent to defining it using the set of technologically feasible allocations, as in Debreu (1951)'s coefficient of resource utilization.

[12] Our measure of aggregate productivity can be used to order feasible sets under a given status quo. The ordering of two feasible sets may flip for different status quos, similar to Scitovszky (1941). However, since the status quo is the initial equilibrium, the choice of status quo is not arbitrary and is disciplined by the data.

2.3 Other Aggregate Measures

For comparison, we define some other common measures of aggregate efficiency.

Chain-weighted Real GDP. In national income accounting, real GDP is measured using approximations to the Divisia (1925) index. Real GDP, defined using the Divisia index, is

$$\log Y(t) = \int_0^t \sum_i \frac{p_i(s)c_i(s)}{\sum_{i'} p_{i'}(s)c_{i'}(s)} \frac{d \log c_i(s)}{ds} ds,$$

where $c_i(s) = \sum_{h \in H} c_{hi}(s)$ denotes aggregate consumption of good i at $s \in [0, t]$.

Kaldor-Hicks (or Cost-Benefit Efficiency, or Sum of Willingness-to-Pay). Another popular aggregate measure is the Kaldor-Hicks efficiency measured in monetary terms.[13] This measure compares the sum of compensating incomes to aggregate income at t. If the sum of compensating incomes, the income needed to make the household indifferent to the status quo, is less than aggregate income, then the winners can hypothetically compensate the losers and there can still be money left-over. The amount of money left over is a measure of the increase in efficiency. This method is the foundation of most of cost-benefit style analyses in applied welfare economics and policy evaluation in public finance and industrial organization.

To write this measure formally, let $e_h(\boldsymbol{p}, u_h)$ be an expenditure function representing preferences $\succeq_h$. The Kaldor-Hicks measure of efficiency at t is

$$A^{KH}(t) = \frac{\sum_h e_h(\boldsymbol{p}(t), u_h(t))}{\sum_h e_h(\boldsymbol{p}(t), u_h(0))}. \tag{7}$$

Note that, by construction, Kaldor-Hicks efficiency at the status quo is equal to one: $A(0) = 1$.[14]

[13] Barcons et al. (2026) define a different version of Kaldor-Hicks efficiency using units of a "welfare numeraire" rather than monetary units. Their definition of Kaldor-Hicks efficiency is the sum of marginal utilities in response to a perturbation divided by the marginal utility of perpetual consumption. This definition may not coincide with the one we use in this paper, even in sign, without some additional assumptions. Similarly, the notion of Kaldor-Hicks efficiency used in Barcons et al. (2026) does not coincide with $A(t)$ either.

[14] In cost-benefit analysis, Kaldor-Hicks efficiency is usually written as the sum of compensating variations. To see that (7) is related to the sum of compensating variations, define the compensating variation for agent h as $cv_h(t) = e_h(\boldsymbol{p}(t), u_h(t)) - e_h(\boldsymbol{p}(t), u_h(0))$. Then, we can rewrite $A^{KH}(t)$ as $A^{KH}(t) = 1/(1 - \sum_h cv_h(t)/\sum_h e_h(\boldsymbol{p}(t), u_h(t)))$. That is, $A^{KH}(t)$ is an increasing function of the sum of compensating variations. We write this transformation to ensure that $A^{KH}(0) = 1$ and, as we shall see later, this transformation coincides with $A(t)$ and $Y(t)$ under some strong but widely used assumptions.

Consumption-equivalent of Representative Agent. Another well-known aggregate measure, when a representative agent exists, is the consumption-equivalent variation used by Lucas (1987). A *representative agent* is a hypothetical single consumer such that the demand of the representative agent for each good, given prices and total income, coincides with equilibrium quantity of that good, given the same prices and aggregate income. (For a formal definition, see Appendix B).

If a representative agent exists, define the consumption-equivalent for the representative agent, $A^{RA}(t)$, to be

$$u^{RA}\left(\boldsymbol{c}^{RA}(t)/A^{RA}(t)\right) = u^{RA}\left(\boldsymbol{c}^{RA}(0)\right),$$

where u^{RA} is the utility function of the representative agent. In words, $A^{RA}(t)$ is the amount by which the aggregate consumption bundle in t must be contracted to make the positive representative agent exactly indifferent to the status quo. As with all the other measures, $A^{RA}(0) = 1$ by construction.

3 Characterization via a Representative Agent

Aggregate productivity can be computed directly from Definition 2. Scale the productivity of all factors by a common scalar $1/Z$, and consider the set of equilibria with transfers given Z, productivities $\boldsymbol{z}(t)$, and wedges $\boldsymbol{\mu}(t)$. If there exists an equilibrium consumption allocation in which all agents are strictly better off than in the status quo, increase Z. If in every equilibrium some agent is worse off than in the status quo, decrease Z. Aggregate efficiency $A(t)$ is the largest value of $1/Z$ for which there exists an equilibrium allocation that leaves everyone at least indifferent to the status quo.

In this section, we provide a useful alternative approach to computing $A(t)$. In Section 3.1, we show that, under some assumptions, computing $A(t)$ is formally equivalent to solving for the equilibrium of a fictional representative agent economy. This result is useful because, when it holds, theorems that are true in representative agent economies can be deployed to study $A(t)$. We use this result in Section 3.2 to show that, under some common but strong assumptions, $A(t)$ coincides with real GDP, Kaldor-Hicks efficiency, and the consumption-equivalent of a positive representative agent. This provides a clean benchmark and the rest of the paper is devoted to understanding deviations from it.

3.1 Equilibrium with Compensated Representative Agent

First, define *individual h's consumption-equivalent* variation.[15]

Definition 3 (Consumption-equivalents)**.** Let $u_h(\boldsymbol{c}_h)$ denote a utility representation for agent h. The *individual consumption-equivalent* function $\tilde{u}_h(\boldsymbol{c}_h)$ is implicitly defined by

$$u_h(\frac{\boldsymbol{c}_h}{\tilde{u}_h}) = u_h(\boldsymbol{c}_h(0)).$$

By construction, the consumption-equivalent function, $\tilde{u}_h$, is homogenous of degree one in consumption and equal to 1 at the status quo $\boldsymbol{c}_h(0)$.[16] The magnitude of $\tilde{u}_h(\boldsymbol{c}_h)$ is measures the amount the consumption bundle $\boldsymbol{c}_h$ has to be scaled to make the household exactly indifferent to the status quo.

Example 1 (Single good). Suppose there is a single consumption good, so $u_h(c_h)$ is some increasing function. In this case,

$$\tilde{u}_h(\boldsymbol{c}_h) = \frac{c_h}{c_h(0)},$$

regardless of the functional form of u_h.

We now define equilibrium with a fictional representative agent.

Definition 4 (Equilibrium with Compensated Representative Agent)**.** An *equilibrium with a compensated representative agent* is the general equilibrium of an economy with the same technologies, resource constraints, and wedges as the original economy but where there is a representative agent with preferences

$$U(\boldsymbol{c}) = \min_h \{\tilde{u}_h(\boldsymbol{c}_h)\}.$$

Note that $U(\boldsymbol{c})$ depends on the status quo allocation because $\tilde{u}_h(\boldsymbol{c}_h)$ is a consumption-equivalent relative to the status quo allocation.[17] The equilibrium with the compensated representative agent is not of direct interest, but is instead a useful device to calculate changes in aggregate productivity. For any equilibrium variable in the decentralized

[15]In macroeconomics, this object is most closely associated with Lucas (1987). However, in the literature on duality in optimization, this function is also called known as the *distance* (see, for example, Cornes, 1992).

[16]The preference relation $\succeq_h$ is homothetic, if and only if, $\tilde{u}$ is a cardinalization of $\succeq_h$. See Appendix D for a discussion of the relationship between $\tilde{u}_h$ and $\succeq_h$ when preferences are non-homothetic.

[17]We call this agent the compensated representative agent because, as we show in Appendix A, the budget shares generated by these preferences are the average of the compensated budget shares of all the agents weighted by each agent's compensating income.

equilibrium X, denote that same variable in the equilibrium with the compensated representative agent by X^{comp}.

Theorem 1 ($A(t)$ via Compensated Representative Agent)**.** *For $t \geq 0$, suppose:*

(i) there exists $\boldsymbol{c}^*(t) \in \mathcal{C}\big(t, 1/A(t)\big)$ such that $u_h\big(\boldsymbol{c}_h^*(t)\big) = u_h(c_h(0))$ for all h;

(ii) wedges $\boldsymbol{\mu}(t)$ are invariant to Z in the equilibrium with a compensated representative agent.

Then the following is true:

$$A(t) = U(\boldsymbol{c}^{comp}(t)) = Y^{comp}(t) = A^{KH,comp}(t).$$

Moreover, at the status quo $t = 0$, prices and quantities in the equilibrium with the compensated representative agent coincide with those in the decentralized equilibrium and $A(0) = 1$.

Theorem 1 shows that if conditions (i) and (ii) hold, then solving for aggregate productivity $A(t)$ is equivalent to solving a representative-agent equilibrium given productivity parameters $\boldsymbol{z}(t)$ and wedges $\boldsymbol{\mu}(t)$. Utility in that representative-agent economy equals $A(t)$ — and since the representative agent has homothetic preferences — this also coincides with real GDP and Kaldor-Hicks efficiency. Furthermore, the consumption allocation $\boldsymbol{c}^{\text{comp}}(t)$ in the equilibrium with the compensated representative agent is interpretable. As we show in the proof, $\boldsymbol{c}^{\text{comp}}(t) = A(t)\boldsymbol{c}^*(t)$ — that is, $\boldsymbol{c}^{\text{comp}}(t)$ is a feasible consumption allocation such that every agent can be made indifferent to the status quo with a fraction $1 - 1/A(t)$ of every good left over. At the status quo technologies and wedges, $A(0) = 1$ and the equilibrium with the compensated representative agent replicates the status quo.

Conditions (i) and (ii) are relatively mild. We discuss them in turn. Condition (i) states that every agent can be made exactly indifferent to the status quo in (6). This is satisfied if, in equilibrium, every agent's utility changes in response to a lump-sum transfer. Intuitively, if an agent is strictly better off in $\boldsymbol{c}^*(t)$ than the status quo, then that agent can be taxed and the proceeds distributed to other agents, allowing a greater reduction in Z. Hence, violations of (i) require cases where such transfers are not feasible. For example, if some agent h is in autarky from the rest of the agents in H, then the only decentralized equilibria feature no lump-sum tax on that agent. In this case, condition (i) is violated and Theorem 1 cannot be used. However, outside of such pathological examples, condition (i) is likely to be satisfied.

Condition (ii) requires that, in the equilibrium with the compensated representative agent, scaling the productivity of factor endowments does not alter the wedges. Since preferences and technologies in the equilibrium with the compensated representative agent are both constant-returns-to-scale, this assumption is likely to hold for many models of endogenous wedges. Of course, the condition is also trivially satisfied if the wedges are exogenously given, or if the equilibria at t is perfectly competitive: $\boldsymbol{\mu}(t) = \mathbf{1}$. This latter case is of importance for measuring misallocation (or the distance of the status quo from the Pareto efficient frontier).

If conditions (i) and (ii) do not hold, then $A(t)$ is still well-defined, but cannot be computed using Theorem 1. However, we maintain these assumptions for the rest of the paper, except in Section 6, where we abandon lump-sum transfers. In that section, we generalize the definition of $A(t)$ to allow for restrictions on redistributive tools. In that case, condition (i) can fail even in non-pathological examples, if there are not enough redistributive tools available to compensate losers.

Theorem 1 is expressed in terms of endogenous variables in the equilibrium with the compensated representative agent, which is a well-understood problem. For this reason, we present the full characterization of variables in that equilibrium in Appendix E. Before using Theorem 1 to construct heterogeneous-agent generalizations of well-known results, we first point out an important, but highly restrictive, special case, where our measure of aggregate productivity coincides with the other popular alternatives.

3.2 A Miraculous Consensus

A widely used but restrictive assumption is that every household has identical homothetic preferences and all households face the same relative prices (i.e. there are no household-specific wedges as in models with incomplete markets). Under these assumptions, every road to defining a measure of aggregate economic activity leads to the same answer.

Proposition 1 (Miraculous Consensus)**.** *If households have identical homothetic preferences, and face the same relative prices, then a positive representative agent exists and*

$$A(t) = Y(t) = A^{KH}(t) = A^{RA}(t).$$

In words, the change in aggregate productivity matches the change in chain-weighted index of real GDP, Kaldor-Hicks (cost-benefit) efficiency, and the consumption-equivalent of the positive representative agent all in the decentralized equilibrium. Hence, under these assumptions, one can compute $A(t)$ without relying on the compensated representative agent.

In the rest of the paper, we focus on cases where consensus does not hold between $A(t)$ and the rest. In these cases, we show through examples that the other measures have undesirable and paradoxical properties. In Section 4, we consider non-identical preferences, and in Section 5, we allow for different households to pay different relative prices for the same goods (nesting incomplete market models). In Section 6, we generalize the definition of $A(t)$ to allow for limits on redistributive tools, which also can cause Proposition 1 to break down.[18]

4 Perfectly Competitive Economies

In this section, we characterize how $A(t)$ responds to changes in technologies, $\boldsymbol{z}(t)$, in perfectly competitive economies where all wedges are equal to one. Section 4.1 provides a version of Hulten's theorem to analyze aggregate productivity with heterogeneous agents. Section 4.2 shows that, even in perfectly competitive economies, once there is nontrivial heterogeneity in agents' preferences, real GDP and Kaldor-Hicks display pathological properties that $A(t)$ does not. Section 4.3 provides some analytical examples to give intuition about $A(t)$ including a generalization of Arkolakis et al. (2012) to an environment with heterogeneous households with non-homothetic preferences.

4.1 Hulten's Theorem with Heterogeneous Agents

Denote the *Domar* weight of each producer or factor i by

$$\lambda_i(t) = \frac{p_i(t)y_i(t)}{\sum_{i'} p_i(t)c_i(t)} \mathbf{1}\{i \text{ is a producer}\} + \frac{w_i(t)z_i(t)L_i}{\sum_{i'} p_i(t)c_i(t)} \mathbf{1}\{i \text{ is a factor}\}.$$

This the sales of i divided by total final expenditures. Recall that for factor i, the quantity of the factor is z_i. The following is a well-known result characterizing changes in perfectly competitive economies.

Proposition 2 (Hulten's Theorem)**.** *The change in chain-weighted real GDP is*

$$\log Y(t) = \int_0^t \sum_i \lambda_i(s) \frac{d \log z_i}{ds} ds. \tag{8}$$

[18] In this paper, we are focused on household heterogeneity, but even with a single household, the miraculous consensus breaks down when preferences are non-homothetic. This point is discussed in detail by Baqaee and Burstein (2023). Intuitively, when preferences are non-homothetic, even for a single agent, scaling the production possibility set ($A(t)$), the budget constraint ($A^{KH}(t)$), and the equilibrium consumption allocation ($A^{RA}(t)$) do not coincide with one another since, as we shrink resources, the household would want to change the bundle of goods they consume.

The most well-known consequence of this result, given by differentiating with respect to t, is that $d\log Y/dt = \sum_i \lambda_i d\log z_i/dt$. This formula, which generalizes Solow (1957), shows that the elasticity of real GDP to the productivity of producer i or the quantity of factor i is just the Domar weight of i.

Theorem 1 immediately implies the following version of Hulten's theorem for $A(t)$.

Proposition 3 (Compensated Hulten's Theorem). *The change in aggregate efficiency at t is*

$$\log A(t) = \int_0^t \sum_i \lambda_i^{comp}(s) \frac{d\log z_i}{ds} ds. \tag{9}$$

In Appendix E, we characterize $\lambda_i^{comp}(s)$ explicitly as a function of the productivity changes $\Delta\log \boldsymbol{z}$, elasticities of substitution, and expenditure shares.

Differentiating (9) with respect to t and evaluating at $t = 0$ shows that, to a first-order approximation, the change in aggregate efficiency, $\Delta\log A$, coincides with the change in real GDP in the competitive equilibrium $\Delta\log Y$.

Corollary 1 (First Order Changes in Aggregate Productivity). *To a first-order approximation, the change in aggregate efficiency is*

$$\Delta\log A \approx \sum_i \lambda_i^{comp}(0)\Delta\log z_i = \sum_i \lambda_i(0)\Delta\log z_i \approx \Delta\log Y \approx \Delta\log A^{KH}.$$

In words, the first-order version of Hulten's theorem applies unaltered to $A(t)$. The final equality, which is standard, shows that real GDP is also equal to Kaldor-Hicks efficiency to a first-order approximation. Hence the miraculous consensus of Theorem 1 holds to a first-order approximation in perfectly competitive economies, even if there are heterogeneous agents with non-homothetic preferences.

Baqaee and Farhi (2019c) show that, to a second-order approximation, changes in real GDP are given by

$$\Delta\log Y \approx \sum_i \lambda_i \Delta\log z_i + \frac{1}{2}\sum_i \Delta\lambda_i\Delta\log z_i.$$

Differentiating (9) twice with respect to t and evaluating at $t = 0$, gives the following extension of Baqaee and Farhi (2019c) to multi-agent settings.

Corollary 2 (Second Order Changes in Aggregate Productivity). *To a second-order approximation, the change in aggregate efficiency due to changes in primitives is*

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

where λ_i and $\Delta\lambda_i^{comp}$ are evaluated at status quo. In Appendix E, we write $\Delta\lambda_i^{comp}$ explicitly as a function of the productivity changes $\Delta \log z$, microeconomic elasticities of substitution, and expenditure shares in the status quo.

Corollary 2 shows that, if the economy is efficient, then discrepancies between aggregate productivity $\Delta \log A$ and real GDP $\Delta \log Y$ start at the second-order, since, generically $\Delta\lambda \neq \Delta\lambda^{\text{comp}}$ (explicit formulas are in Appendix E).[19] The gap between $\Delta\lambda_i$ and $\Delta\lambda_i^{\text{comp}}$ arises because the compensated representative agent's demand responds differently to shocks than aggregate demand in the decentralized economy. The compensated representative agent's demand is constructed to preserve indifference with status quo for all households, whereas the decentralized aggregate demand comes from utility maximization by households whose incomes evolve according to equilibrium changes in relative factor prices.

4.2 Some Paradoxes of Real GDP and Kaldor-Hicks

We now contrast $A(t)$ with real GDP, $Y(t)$, and Kaldor-Hicks efficiency $A^{KH}(t)$, showing that when these measures disagree with one another, the latter can behave pathologically. In particular, when the conditions of the miraculous consensus fail, real GDP and Kaldor-Hicks efficiency do not perform the tasks they were designed to perform. In particular, real GDP can decline even though the economy is producing more of every single good, and Kaldor-Hicks efficiency can rise even though winners cannot compensate the losers.

Real GDP. It is well-known that Divisia-based indices, like real GDP, suffer from paradoxes unless households have identical and homothetic preferences (see Hulten, 1973). Specifically, if households do not have identical and homothetic preferences, then generically, the value of real output $Y(t)$ can be any positive number, regardless of the technology parameters $\boldsymbol{z}(t)$ depending on the path of integration. In particular, every element of $\boldsymbol{c}(t)$ can be lower than $\boldsymbol{c}(0)$ and yet, $Y(t)$ can be greater than $Y(0)$ (see Appendix D for a worked-out Cobb-Douglas example).[20] Hence, although $A(t)$ and $Y(t)$ coincide up to a first-order approximation at $t = 0$, they are not the same nonlinearly.

[19]To derive an expression for $\Delta\lambda^{\text{comp}}$ in terms of microeconomic primitives, we use the fact that $\Delta\lambda^{\text{comp}}$ is the change in Domar weights in a special case of the environment considered by Baqaee and Farhi (2019c) where the consumption growth of each agent is treated as-if it is a final good, and there is a Leontief final demand aggregator over final goods.

[20]The technical reason is the following. Real GDP $\log Y(t)$ is a line integral, and unless preferences are identical and homothetic, the vector field defined by Domar weights is not conservative, making $\log Y(t)$ dependent on the path of integration. See Baqaee and Burstein (2023) for related discussions.

Kaldor-Hicks Efficiency. The following proposition shows that $A(t)$ and $A^{KH}(t)$ are not globally the same. Furthermore, $A^{KH}(t) > 1$ even though it is not feasible for the winners to compensate the losers.

Proposition 4 (Paradox for Kaldor-Hicks Efficiency)**.** *For any change in technologies (movements of the Pareto efficient frontier), we have:*

$$\Delta \log A \leq \Delta \log A^{KH}.$$

For pure redistributions (movements along the Pareto efficient frontier), the change in $A(t)$ is zero, but the change in Kaldor-Hicks efficiency can be positive:

$$\Delta \log A = 0 \leq \Delta \log A^{KH}.$$

These inequalities are strict outside of knife-edge cases. The final inequality is a restatement of the Boadway (1974) paradox. It states that $A^{KH}(t)$ assigns strictly positive value to pure redistributions when relative prices respond to transfers.[21]

Figure 1 illustrates the Boadway paradox using a two-good, two-consumer economy. Intuitively, redistributions lower the relative price of those goods that are more valued by the losers. Hence, in the new equilibrium, it is relatively cheap to compensate these households. Of course, such compensations are, in practice, infeasible because if they were to occur, then relative prices would rise for those households that need compensation. Graphically, $\log A^{KH} > 0$ because $\boldsymbol{c}(0)$ is affordable given prices $\boldsymbol{p}(t)$ (the dashed line is the aggregate budget constraint given new equilibrium prices). In contrast, $\log A = 0$, because $\boldsymbol{c}(t)$ and $\boldsymbol{c}(0)$ are on the same utility possibility frontier.

By continuity, it is possible to construct examples where $\log A > 0$ and $\log A^{KH} < 0$. To do so, we can pair a pure redistribution, which raises A^{KH} but does not affect A with a negative productivity shock, which lowers A. If we pick the size of the redistribution and technology shock appropriately, we can cause A^{KH} to rise while A falls.[22]

Of course, there is another important reason (besides endogeneity of prices) why our measure of efficiency can differ from the Kaldor-Hicks measure. The Kaldor-Hicks measure, by summing up compensating variations, implicitly assumes that lump-sum transfers are available, so that winners can costlessly compensate the losers (assuming relative

[21] See also Blackorby and Donaldson (1990) for a related critique of the sum of compensating variations as a measure of efficiency. See also Jones (2002) for a detailed discussion.

[22] There are some special cases where $A^{KH}(t)$ and $A(t)$ coincide globally in perfectly competitive economies. This happens if relative prices do not depend on final demand, (e.g. as in models with only one primary factor). See Proposition 10 in Appendix C for a formal statement.

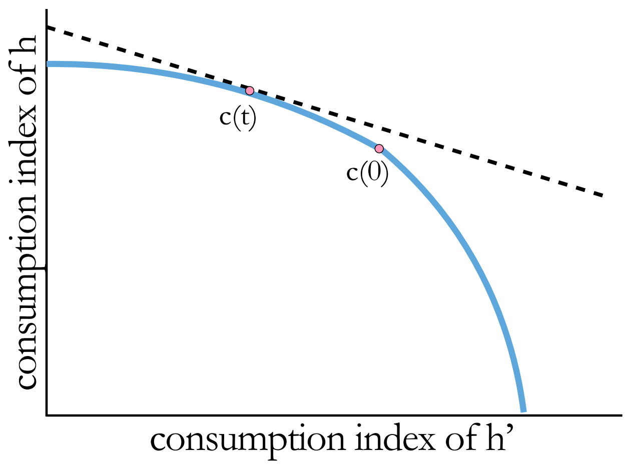


Figure 1: Redistribution from h' to h. The sum of compensating variations at $\boldsymbol{p}(t)$ is greater than zero because the status quo allocation $\boldsymbol{c}(0)$ is below the dashed straight line.

prices are constant). By contrast, our definition of aggregate productivity naturally extends to allow for limited redistribution, as we discuss further in Section 6.

4.3 Analytical Examples

To build some intuition, we work through some analytical examples of how aggregate productivity responds to microeconomic productivity shocks. Appendix D provides more detailed derivations.

Example 2 (Regional Productivity Shocks). Consider households in different regions, indexed by h, with preferences over tradeable goods and locally produced nontradeable services:

$$u_h(\boldsymbol{c}_h) = c_{hg}^{\alpha} c_{hs}^{1-\alpha}, \qquad \sum_h c_{hg} = z_g, \qquad c_{hs} = z_{hs}.$$

The first equation shows that utility in each region depends on goods and services with the expenditure share on goods equal to α. The second equation is the resource constraint for goods, which clear at the aggregate level, since goods are traded. The third equation is the resource constraint for services, which clear region-by-region, since services are not traded. The parameters z_g and z_{hs} control the endowments of goods and services.

Households in region h own the local endowment of services and own a share χ_h of the aggregate endowment of the traded good. This implies that, in equilibrium, χ_h is the expenditures of each household as a share of total consumption expenditures. The

Domar weight on goods is $\lambda_g = \sum_h \chi_h \alpha = \alpha$ and the Domar weight on services in region h is $\lambda_{hs} = \chi_h(1-\alpha)$. Hence, Domar weights are constant in response to productivity changes.

Before considering the change in aggregate productivity $\Delta \log A$, we start by considering the response of real GDP. By Proposition 2, the change in real GDP is the Domar-weighted sum of productivity shocks:

$$\Delta \log Y = \alpha \Delta \log z_g + (1-\alpha)\mathbb{E}_\chi \left[\Delta \log z_s\right],$$

where $\mathbb{E}_\chi \left[\Delta \log z_s\right]$ is the average productivity shock to services weighted by the vector $\boldsymbol{\chi}$. This expression is exact because in the decentralized equilibrium Domar weights do not change ($\Delta\lambda = 0$). Furthermore, since the Domar weights are constant in the competitive equilibrium, there exists a positive representative agent with Cobb-Douglas preferences over goods and services in all regions with preferences:

$$u^{RA}(\boldsymbol{c}) = c_g^{\alpha} \prod_h c_{hs}^{\chi_h(1-\alpha)}.$$

Since the positive representative agent has homothetic preferences, we also have that $\Delta \log Y = \Delta \log A^{RA}$.

In contrast, by Corollary 2, the change in aggregate productivity, to a second-order, is

$$\Delta \log A \approx \alpha \Delta \log z_g + (1-\alpha)\mathbb{E}_\chi \left[\Delta \log z_s\right] - \frac{1}{2}\frac{(1-\alpha)^2}{\alpha} Var_\chi \left[\Delta \log z_s\right] \leq \Delta \log Y = \Delta \log A^{RA}.$$

As implied by Corollary 1, the first two summands (which are first-order) agree with $\Delta \log Y$. However, $\Delta \log A$ features a nonzero second order term, which is absent from $\Delta \log Y$. The miraculous consensus of Proposition 1 fails because the agents do not have the same preferences. The second-order approximation shows that $\Delta \log A$ is a concave envelope of $\Delta \log Y$ around the status quo $\Delta \log \boldsymbol{z} = 0$ — amplifying negative shocks and mitigating positive shocks to services relative to real output. Intuitively, negative shocks to services are more costly because households in that region can only be compensated using goods which are an imperfect substitute for the loss of services.

Our next example uses Theorem 1 to apply a version of the popular Arkolakis et al. (2012) (ACR) formula to economies with heterogeneous (and non-homothetic) preferences.

Example 3 (Gains from Trade with Heterogeneous Preferences). Consider a country with different consumers, indexed by h, that value domestic consumption c_{hd} and consumption of the foreign good c_{hf}:

$$u_h(\boldsymbol{c}_h) = \left[(\alpha_h)^{\frac{1}{\theta_h}} (u_h(\boldsymbol{c}_h))^{\zeta_h} c_{hd}^{\frac{\theta_h - 1}{\theta_h}} + (1-\alpha_h)^{\frac{1}{\theta_h}} c_{hf}^{\frac{\theta_h - 1}{\theta_h}} \right]^{\frac{\theta_h}{\theta_h - 1}},$$

The parameter α_h controls home bias, $\theta_h > 1$ is the compensated Armington elasticity, and ζ_h controls non-homotheticity for agent h. The domestic good is produced linearly from a labor endowment and trade is balanced. We consider the gains from trade relative to autarky by raising iceberg trade costs to infinity. The country trades with the rest of the world in the status quo.

Let $s_{hd}(0)$ denote household h's budget share on the domestic good in the status quo and $\chi_h(0)$ the expenditures of h relative to total spending. We consider the reduction in aggregate productivity caused by moving to autarky (e.g. by raising iceberg costs to infinity). Replicating the argument from ACR, but for the compensated representative agent, losses from autarky are

$$\Delta \log A = -\log \mathbb{E}_{\chi(0)} \left[(s_{hd}(0))^{\frac{1}{1-\theta_h}} \right] \leq 0, \tag{10}$$

where the expectation is across households h using status quo expenditures weights χ_h. The loss, $-\Delta \log A$, is the increase in labor productivity needed in autarky to allow every household to be kept indifferent to the status quo using lump-sum transfers. Note that $(s_{hd})^{\frac{1}{1-\theta_h}}$ is the ACR formula for the gains from trade for a single agent. The equation above shows that aggregate efficiency losses are the average of these individual losses weighted by expenditures in the status quo (denoted by χ).[23, 24]

[23]Compare (10) to the representative-agent ACR formula. Suppose that agents have common homothetic CES preferences with Armington elasticity θ. Then the losses from autarky are $\Delta \log A^{RA} = \log \left(\mathbb{E}_{\chi} [s_{hd}]^{\frac{1}{1-\theta}} \right)$. If the Armington elasticities are the same, $\theta_h = \theta$, then the losses are larger with heterogeneous agents due to Jensen's inequality. For intuition, consider the case where some household h consumes no home goods, i.e., $s_{hd} = 0$ for some h. In this case, $\Delta \log A = -\infty < \Delta \log A^{RA}$ because it is impossible to compensate h in autarky.

[24]Interestingly, the non-homotheticity plays no role in the sense that the formula does not depend on ζ_h (see Appendix D for a worked-out derivation). For example, if there is a single agent with non-homothetic CES preferences, the ACR formula holds without change as long as we use the compensated trade elasticity. If preferences are non-homothetic, then there is a distinction between the compensated and uncompensated trade elasticities. If we have estimates of the latter, one must use Slutsky's equation to first convert them into the compensated elasticities (see, e.g. Auer et al., 2024).

To get more intuition, consider a second-order approximation of $\Delta \log A$:[25]

$$\Delta \log A \approx \mathbb{E}_{\chi(0)}\left[\frac{\log s_{hd}(0)}{\theta_h - 1}\right] - \frac{1}{2} Var_{\chi(0)}\left[\frac{\log s_{hd}(0)}{\theta_h - 1}\right]. \tag{11}$$

The first term is just an "average" version of the ACR formula in logs — the log ACR formula is applied household-by-household and then averaged using households' share of aggregate income χ_h. The expression is weighted by household expenditures because it is more costly to compensate rich households that experience a reduction in welfare than poor households. The second term is the Jensen's inequality term, and it raises aggregate losses if there is any heterogeneity in households' exposure to traded goods either due to variance in expenditure shares, s_{hd}, or trade elasticities, θ_h. In this sense, heterogeneity raises the costs of autarky since some households are more exposed to trade than average. Once again, since domestic and foreign goods are imperfect substitutes, heterogeneity makes compensating the worst-off households more expensive since those households can only be compensated via domestic goods and the marginal value of domestic goods declines as households are given more domestic goods relative to foreign goods.

---

5 Distorted Economies

We now relax the assumption in the previous section that there are no wedges. In Section 5.1, we characterize how aggregate productivity responds to counterfactual changes in microeconomic technologies and wedges. In Section 5.2, we focus on a specific counterfactual: the efficiency losses due to misallocation as measured by the distance from the Pareto efficient frontier. We develop a generalization of Harberger triangles to measure misallocation in general equilibrium with heterogeneous households.

5.1 Comparative Statics for Changes in Technologies and Wedges

Theorem 1 means that we can convert results about real GDP into results about $A(t)$ applying them to variables in the equilibrium with the compensated representative agent. For example, consider the following generalization of Petrin and Levinsohn (2012).

Proposition 5 (Changes in Aggregate Productivity with Wedges)**.** *In response to changes in*

[25] This is an approximation in $\frac{\log s_{hd}}{\theta_h - 1}$ around $s_{hd} = 1$. To derive this, we follow the strategy in Theorem 3 of Baqaee and Farhi (2019a) who consider the gains from trade with a homothetic representative agent.

wedges and productivities, the change in aggregate efficiency is

$$\Delta \log A = \int_0^t \sum_i \lambda_i^{comp}(s) \left[\left(1 - \frac{1}{\mu_i^{comp}(s)} \right) \frac{d \log y_i^{comp}}{ds} + \frac{1}{\mu_i^{comp}(s)} \frac{d \log z_i}{ds} \right] ds.$$

In Appendix E, we characterize $\lambda_i^{comp}(s)$ and $d \log y_i^{comp}/ds$ explicitly as a function of the productivity changes $\Delta \log z$, elasticities of substitution, and expenditure shares using the results in Baqaee and Farhi (2020).

The compensated Domar weights, λ^{comp}, and quantities, $d \log y^{\text{comp}}$, can be computed using standard methods for inefficient economies with homothetic representative agents (see Appendix E).

We contrast Proposition 5 with Harberger's social welfare formula. In his classic paper, Harberger (1971) argued that the welfare effect of a policy that changes wedges can be computed as

$$\Delta \log Y(t) = \int_0^t \sum_i [p_i(s) - mc_i(s)] \frac{dy_i}{ds} ds \ = \ \int_0^t \sum_i \lambda_i(s) \left(1 - \frac{1}{\mu_i(s)} \right) \frac{d \log y_i}{ds} ds, \quad (12)$$

where the equality uses the fact that final expenditure is the numeraire ($\sum_i p_i(s) c_i(s) = 1$ for every s). In words, he argued that whenever a good's marginal benefit, $p_i(s)$, exceeds its marginal cost, $mc_i(s)$, then expanding its quantity (holding others fixed) raises aggregate output.

We see from these equations that if, say, $\Delta \log y_i \neq \Delta \log y_i^{\text{comp}}$, then real GDP and $A(t)$ differ to a first-order in inefficient economies. (Note that, if agents have the same preferences and face the same relative prices, then $\Delta \log y_i = \Delta \log y_i^{\text{comp}}$ and real output coincides with A consistent with the miraculous consensus in Proposition 1).

The following example illustrates the first-order difference between aggregate efficiency, as measured by $A(t)$, and real GDP/Kaldor-Hicks in a simple example. We show that if prices are not equal to marginal costs, then real GDP and Kaldor-Hicks efficiency can rise, to a first-order, in response to pure redistribution, even when the status quo allocation is Pareto efficient and aggregate productivity is constant. Once again, real GDP does not accurately measure the increase in production and Kaldor-Hicks does not correctly detect potential Pareto improvements.

Example 4 (First-Order Aggregate Productivity vs. Real GDP/Kaldor-Hicks). Consider an economy with two agents, 1 and 2, each buying a single consumption good, c_1 and c_2, produced linearly from a unit endowment of labor and sold at markup μ_1 and μ_2. A lump-sum transfer from 2 to 1 raises the spending on good 1 and lowers the spending on good 2. Holding constant technologies, aggregate productivity, measured using $A(t)$, is unchanged. This is because wedges in this economy do not move the allocation off the Pareto-frontier — instead, they redistribute resources from agent 1 to agent 2 given their incomes.

Using Equation (12), it is easy to show that the effect on real GDP (and also Kaldor-Hicks efficiency) from a pure transfer is, to a first-order,

$$\Delta \log Y \approx \Delta \log A^{KH} \approx \bar{\mu} \left[\frac{\mu_1 - \mu_2}{\mu_1 \mu_2} \right] \Delta \lambda_1, \tag{13}$$

where $\bar{\mu}$ is the harmonic sales-weighted average of markups and $\Delta\lambda_1$ is the change in the share of income (including transfers) earned by household 1. Hence, a pure redistribution from agent 2 to agent 1 raises real GDP and Kaldor-Hicks efficiency if the markup paid by agent 1 is higher than the one paid by agent 2. However, there is no way to compensate agent 1 using lump-sum transfers without going back to the status quo equilibrium (which means there is nothing left-over for agent 2). Since the status quo allocation of this economy is Pareto-efficient, any gains to agent 1 come at the cost of agent 2. Accordingly, and in contrast to real GDP and Kaldor-Hicks, $\Delta \log A$ in this example is equal to zero.

We can push this example even farther: suppose again that the transfer occurs at the same time as a decline in the productivity of labor by $\Delta \log z < 0$. In this case, aggregate productivity falls $\Delta \log A = \Delta \log z$. However, the change in real GDP and Kaldor-Hicks efficiency is given by (13) plus $\Delta \log z$, which can be either positive or negative. That is, it is possible that real GDP and Kaldor-Hicks efficiency assign a strictly higher number to an economy with a strictly smaller consumption possibility frontier, even to a first-order.

We now turn our attention to using $A(t)$ to measure the waste caused by distortions.

5.2 Misallocation and the Distance to Pareto Frontier

We now focus on a particular counterfactual: let the status quo equilibrium feature wedges $\boldsymbol{\mu}(0) \neq \mathbf{1}$ and apply Proposition 5 to compute the economic waste caused by distortions. We do this by eliminating wedges at t, $\boldsymbol{\mu}(t) = \mathbf{1}$, and computing $A(t)$. This quantifies waste in terms of factor endowments — the fraction of factors that can be saved in the absence of wedges while keeping everyone indifferent to the status quo. Equivalently, this is also the fraction of every good that can be saved when all distortions are eliminated.

With a complete structural model, $A(t)$ can be computed using Proposition 5. However, below we derive an approximation that is more intuitive and requires less information to be applied.

Proposition 6 (Harberger Triangles). *To a second-order approximation in* $\log \mu$*, the change in aggregate productivity is*

$$\Delta \log A \approx -\frac{1}{2} \sum_i \lambda_i d \log y_i^{comp} \log \mu_i, \tag{14}$$

where $d \log y_i^{comp} \equiv \sum_j \frac{\partial \log y_i^{comp}}{\partial \log \mu_j} \log \mu_j$ *is the change in the quantity of i caused by the wedges in the compensated equilibrium. The approximation error is order* $\log \mu^3$*. The derivatives and expenditure shares in* (14) *are evaluated at the status quo.*[26] *We provide an explicit formula for* $d \log y_i^{comp}$ *in terms of microeconomic primitives in Appendix E.*

This proposition generalizes deadweight loss triangles to measure aggregate productivity losses from wedges in general equilibrium economies with heterogeneous agents. The proof relies on translating results from Baqaee and Farhi (2024) using Theorem 1.

There are two advantages to using Proposition 6 over and above simply applying Proposition 5 using a fully-spelled out structural model. First, the Harberger triangles formula can be used to get analytical intuition for misallocation costs through the use of loglinearized expressions, as demonstrated below. Second, it is possible to populate the terms in (6) with considerably fewer assumptions about the primitives of the economy — e.g. the drivers of distortions, productivity processes, and so on.

The intuition for (14) is familiar — a wedge on *i* is more costly the higher is the Domar weight and the more elastic is the quantity of *i* relative to the wedge. However, compared to a representative agent model with homothetic preferences, the relevant notion of elasticity here is the one in the equilibrium with the compensated representative agent, not the decentralized one.

We provide some pen-and-paper examples to build intuition.

Example 5 (Misallocation when Markups Vary by Household). Consider the misallocation problem studied by Hsieh and Klenow (2009), but suppose there are multiple agents. Each agent *h* has CES preferences over consumption goods with elasticity of substitution θ_h. We consider a situation in which each agent pays potentially a different markup μ_{hi} on each good *i*.[27] Suppose that all consumption goods are ultimately produced linearly from a single common primary factor called labor, which is inelastically supplied.

[26] Usually, such quadratic expansions are evaluated at the undistorted point. However, since λ_i and $d \log y_i^{\text{comp}}$ are multiplied by one power of $\log \mu$, evaluating these terms at status quo wedges changes the expression only at the third order. Hence, the stated approximation remains valid to a second-order.

[27] Formally, we only define wedges that vary by goods rather than good-individual. Hence, to allow for

We can apply Proposition 6 to write the aggregate productivity losses, up to a second order approximation as

$$\Delta \log A \approx \frac{1}{2}\mathbb{E}_{\chi}\left[\theta_h Var_{\boldsymbol{b}_h}\left[\log \boldsymbol{\mu}_h|h\right]\right], \tag{15}$$

where the expectation uses the vector of household income shares, $\boldsymbol{\chi}$, and the variance uses household budget shares over goods, $\boldsymbol{b}_h$, as weights, all evlauated at the status quo.[28] The larger is $\Delta \log A$, the greater the losses from markups. If all agents have the same preferences and face the same wedges, then the expectation in (15) disappears, and the equation collapses to the single agent case, equation (19), in Baqaee and Farhi (2020).

In words, the reduction in efficiency caused by the markups depends on the average variance in markups paid by each household multiplied by that household's elasticity of substitution. Intuitively, if θ_h is very high, then dispersion in markups faced by h causes a greater reduction in aggregate efficiency. Furthermore, aggregate efficiency falls by more if richer households (those with higher χ_h) are exposed to more markup dispersion. This is because any benefits to richer households from eliminating markup dispersion can be used to compensate other agents. Importantly, this expression does not depend on the average markup paid by each household. A proportional scaling of all markups paid by household h would leave this expression unchanged because increasing all markups on a single household is equivalent to a lump-sum tax on that household, and has no effect on aggregate productivity.

The next example applies equation (15) from the previous example to study the efficiency losses due to imperfect insurance.

Example 6 (Misallocation Due to Financial Market Incompleteness). Consider agents with expected utility

$$u_h(\boldsymbol{c}_h) = \sum_s \frac{c_h(s)^{1-1/\theta}}{1-1/\theta}.$$

States of nature, indexed by s, are all equally likely. The coefficient of relative risk aversion is $1/\theta$ (or equivalently, the elasticity of substitution across states is θ).

Each agent h has income $y_h(s) = a_h + \epsilon_h(s)$, where $\epsilon_h(s)$ is an idiosyncratic shock that sums to zero across agents, $\sum_h \epsilon_h(s) = 0$ for every s, with mean zero for each agent $\mathbb{E}[\epsilon_h(s)|h] = 0$. The status quo allocation is financial autarky, so h's consumption in state

this, we introduce intermediaries between each good and each household. Let hi index the intermediary between good i and household h. We assume that this intermediary charges a markup of μ_{hi} on its marginal cost. The intermediary's marginal cost is just the price of good i.

[28] Formally written out, (15) is $= \frac{1}{2}\sum_h \chi_h\theta_h \sum_n b_{hn}\left[\log \mu_{hn} - \sum_{n'} b_{hn'} \log \mu_{hn'}\right]^2$.

s is $c_h^0(s) = y_h(s)$. The aggregate resource constraint for the economy is $\sum_h c_h(s) = \sum_h a_h$, because, by assumption, $\sum_h \epsilon_h(s) = 0$ for every s.

To decentralize this allocation using wedges, suppose that there are complete state-contingent markets with household-by-state wedges $\mu_h(s)$. Household h's budget constraint can be written as

$$\sum_s \mu_h(s) c_h(s) = a_h.$$

The wedges that decentralize the status quo allocation must satisfy

$$\frac{c_h^0(s)}{c_h^0(s')} = \left[\frac{\mu_h(s)}{\mu_h(s')}\right]^{-\theta}.$$

Substituting these wedges into equation (15) implies that the gains from completing financial markets are approximately given by:

$$\Delta \log A \approx \frac{1}{2}\theta \mathbb{E}_\chi \left[Var\left[\log \mu_h(s)|h\right]\right] = \frac{1}{2}\frac{1}{\theta}\mathbb{E}_\chi \left[Var\left[\log c_h(s)|h\right]\right],$$

where χ_h is household h's expected share of consumption in the status quo, equal to $a_h / \sum_{h'} a_{h'}$. This formula disregards inequality due to dispersion in the persistent component of income (dispersion in consumption caused by a_h) because the variance is conditional on household h. Instead, misallocation depends on the average conditional consumption variance, weighted by household income. Holding the consumption process fixed, the gains from completing financial markets are larger, the higher is the risk aversion $1/\theta$ parameter.

The example above is simple, but hints at a much more general idea. Baqaee and Burstein (2025b) build on this basic idea to analyze and quantify the losses from financial market imperfections in both open and closed economies. In that paper, we discuss how to allow for dynamics, labor-leisure choice, capital accumulation, borrowing constraints, and international trade. A running theme in that paper is that the Harberger triangles formula can be computed without strong assumptions on the full structure of the model (e.g. productivity/income process, etc.)

The final example in this section shows that misallocation, as measured by $\Delta \log A$, need not equal to changes in real GDP, $\Delta \log Y$, or changes in the welfare of a representative agent, $\Delta \log A^{RA}$, even in cases where a representative agent exists.

Example 7 (Distance to Frontier: Real GDP and Positive Representative Agent). Suppose each agent h's has CES preferences over consumption goods with elasticity of substitution θ_h. Suppose each agent consumes a different selection of goods but all goods are produced linearly from the labor endowment. The markup on the ith good consumed by household h is denoted by μ_{hi}. Suppose that we eliminate markups, each households' share of income, χ_h, stays constant. Since the distribution of income is constant, there is a positive (and in this case, also normative) representative agent with Cobb-Douglas preferences across each households' consumption bundle (i.e. an agent whose utility is maximized by observed allocations). The change in the welfare of this representative agent, in consumption-equivalent terms, is equal to the change in chain-weighted real GDP, and both are equal to a second-order approximation to

$$\Delta \log Y = \Delta \log A^{RA} \approx \underbrace{\frac{1}{2}\mathbb{E}_{\chi}\left[\theta_h Var_{\boldsymbol{b}_h}\left[\log \boldsymbol{\mu}_h | h\right]\right]}_{\approx \Delta \log A} + \frac{1}{2} Var_{\boldsymbol{\chi}}\left[\mathbb{E}_{\boldsymbol{b}_h}\left[\log \boldsymbol{\mu}_h | h\right]\right],$$

where we use (15). The change in real GDP and the welfare of the representative agent are (weakly) larger than the change in aggregate efficiency. One limiting case of this is where all markup-variation is at the household level. In this case, $\Delta \log A = 0$ because the economy is already on the efficient frontier and eliminating markups is purely redistributive. However, in this example, $\Delta \log Y = \Delta \log A^{RA} > 0$.

6 Aggregate Productivity with Costly Redistribution

In this section, we extend the definition of $A(t)$ to allow for imperfect redistributive tools. This is another advantage of our approach relative to measures based on adding up willingness-to-pay across all households (e.g. Kaldor-Hicks). Intuitively, when we add up willingness-to-pay, we implicitly assume that winners can costlessly compensate losers using unrestricted transfers. In Sections 4 and 5, we illustrated one issue with this approach: monetary compensations can change relative prices so that, in practice, the necessary compensations are infeasible. In this section, we focus on a second issue — unrestricted monetary compensations may be infeasible because the required transfers are unavailable for reasons that are exogenous to our analysis (e.g. politics, information, etc.)

Section 6.1 generalize the definition of $A(t)$ to allow for restricted redistributive tools. Section 6.2 extends Hulten's theorem and Harberger triangles to this environment. Section 6.3 extends the losses from autarky exercise of Arkolakis et al. (2012) to allow for heterogeneous agents and compensations that can only be achieved using distortionary taxes. Finally, we close with a quantitative application in Section 6.4, where we analyze the aggregates effect of the rise of China on U.S. workers, allowing for different redistributive tools and degrees of labor market mobility.

6.1 Definition with Limited Redistributive Tools

Suppose that there are ad valorem taxes $\boldsymbol{\tau}$ and lump-sum taxes and transfers $\boldsymbol{T}$ that can be used to compensate agents. Since lump-sum transfers must be denoted in units of some numeraire, we choose total spending as the numeraire. Assume that the value of taxes and transfers, $(\boldsymbol{\tau}, \boldsymbol{T})$ is restricted to some exogenous set of allowable values $\mathcal{T}$.

Define the feasible consumption set, given restrictions on redistributive tools, to be:

$$\mathcal{C}^{\text{costly}}(t, Z) = \left\{ \boldsymbol{c} \in \mathbb{R}^{H \times N} : \text{ there exist } (\boldsymbol{\tau}, \boldsymbol{T}) \in \mathcal{T} \text{ supporting } \boldsymbol{c} \text{ as equilibrium} \right\}.$$

This is the set of consumption allocations that can be supported as equilibria given exogenous parameter values at t, aggregate factor-augmenting technology shifter Z, and tax-and-transfer values belonging to $\mathcal{T}$. We define aggregate productivity with costly redistribution, $A^{\text{costly}}(t)$, analogously to before except using $\mathcal{C}^{\text{costly}}(t, Z)$ in place of $\mathcal{C}(t, Z)$.

Definition 5 (Aggregate Productivity with Costly Redistribution)**.** *Aggregate productivity* at t, given redistributive tools $\mathcal{T}$, is the maximum contraction of factor-augmenting productivity such that every agent can be kept at least indifferent to the status quo allocation. Formally,

$$A^{\text{costly}}(t) \equiv \max \left\{ Z \in \mathbb{R} : \text{there is } \boldsymbol{c} \in \mathcal{C}^{\text{costly}}(t, 1/Z) \text{ and } \boldsymbol{c}_h \succeq_h \boldsymbol{c}_h(0) \text{ for every } h \right\}. \tag{16}$$

Figure 2 illustrates the relationship between $A^{\text{costly}}(t)$ and $A(t)$ using a simple two agent example. The agents are indexed by h and h', and the status quo allocation, $\boldsymbol{c}(0)$, and decentralized allocation without transfers, $\boldsymbol{c}(t)$, are denoted by red circles. In the decentralized allocation, agent h' is better off and agent h is worse off compared to the status quo. The dashed line indicates the set of feasible consumption allocations t given unrestricted lump-sum taxes, $\mathcal{C}(t, 1)$. The solid blue line shows the same set but given restricted and distorting redistributive tools. The two frontiers touch at the decentralized point, since the decentralized point does not engender any distortionary redistributive taxation. However, the solid blue set is strictly smaller than the dashed line since distortionary taxation limits the set of feasible redistributions. In this example, aggregate productivity is measured by the required contraction in the productivity of factors that causes the consumption possibility to intersect the status quo allocation. In this example, the required contraction of $\mathcal{C}(t, 1)$ is larger than the one for $\mathcal{C}^{\text{costly}}(t, 1)$, and so aggregate productivity gains are smaller with distortionary redistributive tools.

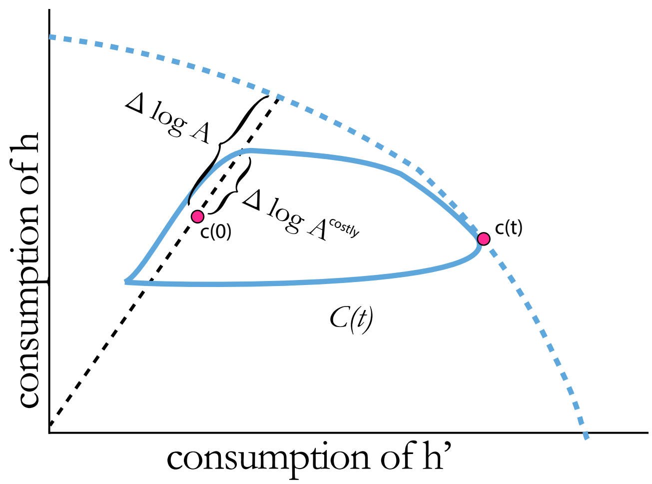


Figure 2: Aggregate productivity with lump-transfers and distortionary taxes.

With limited redistributive tools, Theorem 1 no longer applies, which means that $A^{\text{costly}}(t)$ cannot be calculated through a fictitious representative agent, and must be solved for directly.

6.2 Hulten's Theorem with Costly Redistribution

If there are enough redistributive tools to ensure households can be made exactly indifferent to the status-quo, then starting at a perfectly competitive equilibrium, the change in aggregate productivity still obeys Hulten's theorem even with costly redistribution because the losses from distortionary redistributive taxes are second order.

To formalize this logic, suppose the following condition holds:

$(*)$ there exists $\boldsymbol{c}^*(t) \in \mathcal{C}^{\text{costly}}\left(t, 1/A^{\text{costly}}(t)\right)$ such that $u_h\left(\boldsymbol{c}_h^*(t)\right) = u_h(c_h(0))$ for all h.

This is the same condition as in Theorem 1 and ensures that, given $\boldsymbol{\mu}(t)$, $\boldsymbol{z}(t)$, and $Z = 1/A^{\text{costly}}(t)$, it is possible to keep every agent exactly indifferent to the status quo using the redistributive tools.

Proposition 7 (Hulten's Theorem with Costly Redistribution)**.** *If the status quo is perfectly competitive (no taxes or wedges) and condition $(*)$ holds, then, to a first order approximation in $\Delta \log \boldsymbol{z}$,*

$$\Delta \log A^{costly} = \sum_i \lambda_i(0) \Delta \log z_i.$$

Intuitively, the losses from costly-redistribution are second-order, and hence to a first-order approximation, only the direct effects of the productivity shock matter (assuming we start at a competitive equilibrium).

Of course, the losses from distortionary taxation do matter to higher orders. Indeed, the following proposition shows that, to a second-order, $\Delta \log A^{\text{costly}}$ is lower than $\Delta \log A$ by exactly the deadweight loss triangles caused by distortionary compensations.

Proposition 8 (Productivity Shocks with Limited Redistribution)**.** *If the status quo is perfectly competitive and condition (∗) holds then, to a second order approximation in* $\Delta \log \boldsymbol{z}$*,*

$$\Delta \log A^{costly} = \underbrace{\sum_i \left(\lambda_i + \frac{1}{2} \sum_j \frac{\partial \lambda_i^{comp}}{\partial \log z_j} \Delta \log z_j \right) \Delta \log z_i}_{\Delta \log A} + \frac{1}{2} \sum_i \lambda_i \left(\sum_j \frac{\partial \log y_i^{comp}}{\partial \log \tau_j} \Delta \log \tau_j^* \right) \Delta \log \tau_i^*, \tag{17}$$

where $\boldsymbol{\tau}^*(t)$ *implements the solution in* (16).

The first set of summands are the same as for $\Delta \log A$. The last set of summands, which are new and non-positive, capture the inefficiency caused by imperfect redistribution. These are the sum of Harberger triangles associated with the linear taxes in $\tau^*(t)$. That is, the response of aggregate productivity to technology shocks is the same as it would be if lump-sum transfers were possible minus the deadweight loss triangles associated with distortionary taxes needed to compensate households. The simplicity of Equation (17) follows from the fact that the status quo is undistorted. This ensures that (1) there are no interactions of taxes with pre-existing distortions, (2) the cross-partials between $d \log \tau^*$ and $d \log z$ are all zero due to the envelope theorem.

6.3 Analytical Example: Losses from Autarky

To illustrate Proposition 8, we apply it to quantify the losses from autarky accounting for limited redistribution.

Example 8 (Gains from Trade with Limited Redistribution). We revisit Example 3, which studied the gains from trade, but this time we incorporate limits to redistribution. We add a labor-leisure margin, and assume redistribution can only be done by taxing consumption.

Suppose there are two households, and household h has nested-CES preferences over domestic consumption goods, c_{hd}, imported consumption goods, c_{hf}, and leisure l_h:

$$u_h(\boldsymbol{c}_h) = \left[(1-\gamma_h)^{\frac{1}{\rho}} \left[(1-\alpha_h)^{\frac{1}{\theta}} c_{hd}^{\frac{\theta-1}{\theta}} + \alpha_h^{\frac{1}{\theta}} c_{hf}^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}\frac{\rho-1}{\rho}} + \gamma_h^{\frac{1}{\rho}} l_h^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}.$$

The model in Example 3 did not feature the leisure good. The inner nest combines domestic and foreign consumption goods with Armington elasticity θ and home bias controlled by the parameter α_h. The outer nest combines the goods bundle with leisure with elasticity of substitution ρ and share parameter γ_h. The parameter ρ controls the Frisch elasticity of labor supply.

Household h is endowed with one unit of time and a_h efficiency units of labor and faces a budget constraint:

$$\tau p_d c_{hd} + \tau p_f c_{hf} = w a_h (1 - l_h) + T_h,$$

where p_d and p_f denote the price of each consumption good, w_h is the wage per efficiency unit, τ is the gross-tax rate on consumption, and T_h is a lump sum transfer. Budget balance requires $(\tau - 1)\sum(p_d c_{hd} + p_f c_{hf}) = \sum T_h$. The domestic consumption good is produced linearly with labor, so the resource constraint for domestic consumption is $\sum_h c_{hd} = \sum_h a_h(1 - l_h)$, with $p_d = w$. The resource constraint for leisure is $l_h \leq 1$.

Let

$$\Omega_{hd} = \frac{p_d^0 c_{hd}^0 + p_f^0 c_{hf}^0}{w^0 a_h}$$

denote household h's budget share on consumption in the status quo as a share of the value of h's total time endowment (the remainder is implicit expenditures on leisure). For simplicity of exposition, and since it is fairly realistic, we assume that both households work the same number of hours in the status quo, which implies that $\Omega_{hd} = \Omega_d$ does not vary by household. Let s_{hd} denote household h's share of expenditures on the domestic consumption good relative to all consumption goods:

$$s_{hd} = \frac{p_d^0 c_{hd}^0}{p_d^0 c_{hd}^0 + p_f^0 c_{hf}^0}.$$

The status quo is a competitive equilibrium without taxes in which the country trades with the rest of the world. We consider the loss in productivity from closing the economy to autarky in two cases: (1) where unrestricted lump-sum transfers are available, and (2) lump-sum transfers must be positive and revenues can only be raised using ad valorem consumption taxes.

Lump-Sum Taxation. With lump-sum taxation, using Corollary 2, we can write the losses from autarky to a second-order approximation as

$$\Delta \log A^{\text{lump-sum}} \approx \underbrace{\Omega_d \mathbb{E}_\chi \left[\frac{\log s_h}{\theta - 1}\right]}_{\text{1st order}} \underbrace{- \frac{1}{2}\Omega_d^2 Var_\chi \left(\frac{\log s_h}{\theta - 1}\right) + \frac{1}{2}(\rho - 1)\Omega_d(1 - \Omega_d)\mathbb{E}_\chi \left[\left[\frac{\log s_h}{\theta - 1}\right]^2\right]}_{\text{2nd order with lump-sum taxation}}.$$

This expression is identical to Equation (11) in Example 3 when there is no leisure, $\Omega_d = 1$. The first and second summands are the same as in (11) but are now scaled by Ω_d to account for the fact that households also consume leisure. The final summand, which is absent in (11), accounts for complementarities/substitutabilities between consumption and leisure. If consumption and leisure are complements, $\rho < 1$, then the reduction to consumption caused by autarky reduces the value of leisure through complementarity, raising losses from autarky further.

Linear Taxation. Now consider the case where lump-sum taxation is unavailable so that lump-sum transfers must be non-negative: $\boldsymbol{T} \geq 0$, financed by a uniform consumption tax. Proposition 8 now implies that, to a second-order approximation,

$$\Delta \log A^{\text{costly}} \approx \Delta \log A - \underbrace{\frac{1}{2}\rho\Omega_d(1 - \Omega_d)(d \log \tau^*)^2}_{\text{2nd order losses from distorting taxes}},$$

where τ^* is the consumption tax needed to compensate the household that loses more from autarky. A given tax is more distorting the higher is ρ, which controls substitution between consumption and leisure (ρ can be interpreted as the Frisch elasticity of labor supply), and the closer is Ω_d to $1/2$. If Ω_d is equal to either one (households do not value leisure) or zero (households do not value consumption), then there is no distortion from taxing consumption.

Index the two households by h and h' and suppose that h is more exposed to foreign goods: $s_{hd} < s_{h'd}$. This means that, in the decentralized equilibrium, household h is more negatively affected by autarky than h'. In this case, the optimal compensation raises a consumption tax and sends all collected tax revenues to h. The tax required for the compensation is $d \log \tau^* = \frac{\chi_h}{\theta - 1}\left[\log s_{h'd} - \log s_{hd}\right] > 0$, to a first-order, where χ_h is h's share of aggregate income in the status quo. The required consumption tax is larger the bigger is the heterogeneity in exposure to the trade shock, and the larger is household h's share of aggregate income. Appendix D provides numerical examples and shows that the second order approximation is very accurate even for very open economies.

In Appendix D, we consider another example, where skill-biased technical change affects different workers differently, and show how the second order approximation in Proposition 8 performs.

6.4 Quantitative Example: the China Shock

Our final example quantifies the aggregate productivity gains for U.S. consumers from the rise of China. We use the general equilibrium model in Baqaee and Farhi (2024). In the model, each country has different factor endowments, and we treat owners of different factor endowments as different agents.

The rise of China, which we model by improving technologies in China, changes relative wages amongst the different domestic factors, with different consequences for different agents. We study how the aggregate productivity gains for the U.S. depend on assumptions about factoral mobility (across sectors) and the availability of tax and transfer tools.

Summary of calibration. The model has 9 regions (Canada, China, France, Germany, Great Britain, Japan, Mexico, U.S., and the rest of the world) and 30 industries in each country. Production by each industry is a nested CES aggregator combining four domestic primary factors (low-, medium-, high-skill labor, and capital) with intermediate inputs. The intermediate input bundle used by each industry is a nested CES aggregator over all industries and origin countries. All households in each country consume the same domestic consumption bundle, which is a homothetic nested CES aggregator over all industries and origin countries. (We abstract from heterogeneity in preferences within countries). The initial expenditure shares are calibrated according to the World Input-Output Database in 2008. Since tariffs in 2008 were quite low, the model is calibrated assuming there are no import tariffs.[29]

[29] The elasticity of substitution between primary factors is set to one. The elasticity of substitution between value-added and intermediates is 0.5. Each country-industry pair has a unique bundle of intermediate inputs sourced from different industries with elasticity of substitution across industries of 0.2. Each country also has a unique consumption bundle, with elasticity of substitution across industries of 0.9. Every destination country-industry pair purchases from each industry a unique mix of inputs sourced from different origin countries (e.g. U.S. mining purchases a unique mix of machinery from different origin countries). The Armington trade elasticity is equal to 5.

Shock. In 2008 (the calibration year), China's GDP was roughly 5% of the world's. By 2023, this number had risen to 18%. We model the rise of China through an increase in Chinese factor-augmenting productivity growth (roughly tripling the efficiency units of all Chinese factor endowments) to ensure that China's share of world GDP rises to 18%. We consider how this shock affects the United States under four different scenarios.

Restrictions on redistribution. Each scenario imposes different assumptions on what redistributive tools are available. The scenarios are:

i. *Tariffs with lump-sum transfers:* the government can raise a uniform tariff on imports, and also has access to unrestricted lump-sum transfers (i.e. lump-sum transfers can be positive or negative).

ii. *Tariffs with targeted rebates:* the government can raise a uniform tariff on imports and has full discretion on how to rebate any additional tariff revenues (i.e. if tariff revenues rise after the shock, in units of world GDP, then the government can choose who to rebate that additional revenue to).

iii. *Tariffs with non-targeted rebates:* the government can raise a uniform tariff on all imports but any revenues generated by tariffs are rebated back to domestic households in proportion to their pre-shock initial share of aggregate income.

iv. *Laissez-faire*: there are no redistributive tools available. The consumption possibility set is a single point corresponding to the decentralized equilibrium in the U.S.

In scenarios i., ii. and iii., the consumption possibility set is the set of equilibrium consumption allocations for U.S. consumers given different levels of uniform import tariffs and feasible transfers (taking into account how different tariff levels impact the world equilibrium). It is this set that we expand or contract to ensure all U.S. households can be made at least indifferent to the status quo.[30] The consumption possibility set in each scenario is smaller than the preceding, which implies the productivity gains will be smaller as well.

Status quo. We assume that the 2008 data correspond to laissez-faire in which all taxes are zero. If we were to treat this as the status quo, then even without the China shock aggregate productivity $A > 1$ in scenarios (i), (ii), and (iii), reflecting the gains from

[30] This can be accomplished in general equilibrium by scaling the productivity of all primary factors in the world by some constant. This constant is chosen to ensure that all U.S. consumers can be made at least indifferent to the status quo.

optimal tariffs (since the status quo has no tariffs). Our focus is not on quantifying these tariff gains, but to illustrate how an imperfect redistributive tool — here, an import tariff —influences the aggregate implications of the China shock.

We therefore assume that, if an import tariff is available (scenarios ii, iii, iv), then that import tariff is already set to maximize the quantity of the U.S. consumption good prior to the China shock. This ensures that the status quo allocation is not Pareto dominated by other allocations in the feasible status quo consumption possibility set (i.e. in the absence of the China shock, aggregate productivity is 1 by construction in every scenario). We do this so that we do not conflate the effects of the China shock with the effects of establishing optimal tariffs. Nevertheless, to keep the status quo allocation similar to the data in 2008 — where tariffs were small — we assume that U.S. tariffs provoke symmetric retaliation from the rest of the world. As a result, the optimal U.S. tariff in the status quo is small anyhow (2.3% in the case without factor mobility and 1.0% in the case with factor mobility), ensuring that expenditure shares in the status quo are close to the 2008 data prior to the shock. Appendix F summarizes the computational details.

Results. The aggregate productivity gains for the U.S. due to the China shock is shown in Table 1. We consider two specifications of the model, labelled "immobile" and "mobile" factors. When factors are "immobile", primary factors (labor and capital) in each country-industry pair cannot move across industries. When factors are mobile, there is one national market for each factor type (low-, medium-, high-skill labor and capital) and all industries in a country hire from the national market. Comparing these two specifications reveals the importance of reallocation for determining the aggregate effect of a shock.

Table 1: Effect of China Shock on the United States

Scenario	Immobile Factors		Mobile Factors	
	$\Delta\log A$	Δ Tariff (p.p.)	$\Delta\log A$	Δ Tariff (p.p.)
Tariffs & lump-sum transfers	0.008	+0.1	0.010	+0.3
Tariffs & targeted rebates	-0.010	+7.4	0.010	+0.4
Tariffs & non-targeted rebates	-0.179	+11.4	0.008	+1.6
Laissez-faire	-0.235	–	0.008	–

Tariff changes are expressed in percentage points.

The first row shows the change in aggregate productivity, as measured by A, for the U.S., assuming there are lump-sum transfers. In this case, the miraculous consensus of

Proposition 1 holds, and the response of $\Delta \log A$ coincides with the response of aggregate real consumption in the U.S. This also coincides with real consumption by the representative U.S. household, as well as the change in the Kaldor-Hicks measure of efficiency. In this case, $\Delta \log A$ rises by around 1 log point in response to the China shock.

The second row considers the case where import tariffs are available, but redistribution can only draw on excess tariff revenues. In this case, $\Delta \log A$ falls by 1 log point if factors cannot move across sector. This is because the China shock causes real wages in some sectors to fall. To compensate these households, we must raise tariff revenues by around 7%. These tariffs, which trigger retaliatory tariffs from the rest of the world, cause overall efficiency to decline. In other words, in order to make every U.S. household at least as well off after the China shock requires expanding the consumption possibility set by 1 log point. The picture is very different if factors are mobile, since in this case, real wages do not decline, and hence large tariffs are not needed to offset real income losses suffered by some subset of households.

The third row considers the case where tariff revenues can only be distributed according to pre-shock (initial) income shares. Because redistributive tools are much more severely restricted, compensating losers becomes much harder. It may be impossible to equalize changes in real consumption across consumers, so condition ($*$) need not hold. Accordingly, aggregate productivity in the U.S. falls by 17.9 log points since compensating workers in losing industries requires instituting very large tariffs (around 11.4%), which in turn trigger a large trade war. Once again, these restrictions on redistributive policy are much less important if factors are mobile across sectors, and the overall efficiency gains are barely affected.

The final row considers the laissez-faire case, where $\Delta \log A$ reduces to the change in real consumption for the workers whose real wages decline the most under laissez-faire. When factors are immobile, this is the *Textile and Leather Products* sector, where real wages decline by around 23 log points in response to the China shock. The contrast with the mobile-factor case is stark: with factor mobility, heterogeneity across factors is substantially attenuated, so the change in aggregate productivity closely approximates that under full redistribution.

The results in Table 1 show that the change in aggregate efficiency depends strongly on (a) the extent to which shocks have asymmetric effects on households, and (b) the redistributive tools available. Note that although the redistributive tools are important for quantifying the change in aggregate productivity reported in Table 1, we do not take a stance on how these tools should be used in practice. For example, if lump-sum transfers are available, the consumption possibility set can be contracted by 1% and everyone can be kept at least as well off; hence, after the shock and compensations, there is a 1% surplus in the U.S. consumption good. We measure this surplus without taking a stance on how it should be distributed.

7 Conclusion

We generalize the cost-benefit approach to aggregate efficiency to environments with heterogeneous agents, general equilibrium, and limited redistribution. Our measure, which converts shocks into a welfare-equivalent change in total factor productivity, collapses to the Solow residual and Kaldor-Hicks efficiency when households have the same homothetic preferences and face the same relative prices.

We show how to compute this measure by solving the equilibrium of an economy with some fictional as-if representative agent. This provides a method to translate theorems and tools about representative-agent economies to study aggregate efficiency in economies with heterogeneous agents, including, for example, Hulten (1978), Harberger (1964), Hsieh and Klenow (2009), Arkolakis et al. (2012), Baqaee and Farhi (2019c), and Baqaee and Farhi (2020). This also allows us to calculate aggregate productivity using only information on observables like expenditure shares and price elasticities of supply and demand curves.

We also contrast our measure of aggregate productivity with real GDP, measures of efficiency based on the sum of compensating variations (Kaldor-Hicks efficiency), and consumption-equivalents for positive representative agents (if there is one). We show that these popular alternative measures of efficiency have serious flaws.

In two stand-alone companion papers, we apply the theoretical results of this paper to other contexts where household heterogeneity is central. Baqaee and Burstein (2025b) consider losses in aggregate efficiency from financial market incompleteness, within and across borders. Baqaee and Burstein (2025a) characterize aggregate productivity in random utility models with discrete choice, as in spatial economies, where households make different choices due to differences in preferences. An interesting extension, which we do not pursue in this paper, is to study policy problems where maximizing aggregate productivity is the policymaker's objective.

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012). The network origins of aggregate fluctuations. *Econometrica 80*(5), 1977–2016.

Allais, M. (1979). La théorie générale des surplus.

Antràs, P., D. Chor, T. Fally, and R. Hillberry (2012). Measuring the upstreamness of production and trade flows. *The American Economic Review 102*(3), 412–416.

Antras, P., A. de Gortari, and O. Itskhoki (2017). Globalization, inequality and welfare. *Journal of International Economics 108*(C), 387–412.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012). New trade models, same old gains? *American Economic Review 102*(1), 94–130.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of economic theory 2*(3), 244–263.

Auer, R., A. Burstein, S. Lein, and J. Vogel (2024). Unequal expenditure switching: Evidence from switzerland. *Review of Economic Studies 91*(5), 2572–2603.

Auerbach, A. J. and L. J. Kotlikoff (1987). *Dynamic Fiscal Policy*. Cambridge University Press.

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

Baqaee, D. and E. Farhi (2019a, July). Networks, barriers, and trade. Working Paper 26108, National Bureau of Economic Research.

Baqaee, D. and E. Farhi (2019b). A short note on aggregating productivity. Technical report, National Bureau of Economic Research.

Baqaee, D. and E. Rubbo (2023). Micro propagation and macro aggregation. *Annual Review of Economics 15*(1), 91–123.

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). Efficiency costs of incomplete markets. Working Paper 34233, National Bureau of Economic Research.

Baqaee, D. R. and E. Farhi (2019c). 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.

Baqaee, D. R., E. Farhi, and K. Sangani (2024). The darwinian returns to scale. *Review of*

Economic Studies 91(3), 1373–1405.

Barcons, S., E. Dávila, and A. Schaab (2026). Intergenerational welfare assessments. Technical report, National Bureau of Economic Research.

Bergson, A. (1938, February). A reformulation of certain aspects of welfare economics. *The Quarterly Journal of Economics 52*(2), 310–334.

Bhandari, A., D. Evans, M. Golosov, and T. Sargent (2021). Efficiency, insurance, and redistribution effects of government policies. Technical report, Working paper.

Bigio, S. and J. La'O (2016). Financial frictions in production networks. Technical report.

Blackorby, C. and D. Donaldson (1990). A review article: The case against the use of the sum of compensating variations in cost-benefit analysis. *canadian Journal of Economics*, 471–494.

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

Bornstein, G. and A. Peter (2024). Nonlinear pricing and misallocation. Technical report, National Bureau of Economic Research.

Comin, D., D. Lashkari, and M. Mestieri (2021). Structural change with long-run income and price effects. *Econometrica 89*(1), 311–374.

Cornes, R. (1992). *Duality and modern economics*. CUP Archive.

Dávila, E. and A. Schaab (2022). Welfare assessments with heterogeneous individuals. Technical report, National Bureau of Economic Research.

Dávila, E. and A. Schaab (2023). Welfare accounting. Technical report, National Bureau of Economic Research.

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

Debreu, G. (1954). A classical tax-subsidy problem. *Econometrica: Journal of the Econometric Society*, 14–22.

Dekle, R., J. Eaton, and S. Kortum (2008). Global rebalancing with gravity: measuring the burden of adjustment. *IMF Staff Papers 55*(3), 511–540.

Divisia, F. (1925). L'indice monétaire et la théorie de la monnaie. *Revue d'économie politique 39*(4), 842–861.

Donald, E., M. Fukui, and Y. Miyauchi (2023). Unpacking aggregate welfare in a spatial economy. Technical report, Tech. rep.

Dupuit, A. J. É. J. (1844). De la mesure de lútilité des travaux publics. *Annales des ponts et chaussées 8*.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the royal statistical society series a: statistics in society 120*(3), 253–281.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica 79*(3), 733–772.

Galle, S., A. Rodríguez-Clare, and M. Yi (2023). Slicing the pie: Quantifying the aggregate and distributional effects of trade. *The Review of Economic Studies 90*(1), 331–375.

Harberger, A. C. (1954). Monopoly and resource allocation. In *American Economic Association, Papers and Proceedings*, Volume 44, pp. 77–87.

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.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of political economy 63*(4), 309–321.

Heathcote, J. and H. Tsujiyama (2021). Optimal income taxation: Mirrlees meets ramsey. *Journal of Political Economy 129*(11), 3141–3184.

Hendren, N. and B. Sprung-Keyser (2020). A unified welfare analysis of government policies. *The Quarterly journal of economics 135*(3), 1209–1318.

Hicks, J. R. (1939). The foundations of welfare economics. *The economic journal 49*(196), 696–712.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing TFP in China and India. *The quarterly journal of economics 124*(4), 1403–1448.

Hulten, C. R. (1973). Divisia index numbers. *Econometrica: Journal of the Econometric Society*, 1017–1025.

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

Jones, C. (2002). The" boadway paradox" revisited.

Kaldor, N. (1939). Welfare propositions of economics and interpersonal comparisons of utility. *The economic journal 49*(195), 549–552.

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

Liu, E. (2017). Industrial policies and economic development. Technical report.

Lucas, R. E. (1987). *Models of business cycles*, Volume 26. Basil Blackwell Oxford.

Luenberger, D. G. (1996). Welfare from a benefit viewpoint. *Economic Theory 7*(3), 445–462.

Mas-Colell, A., M. D. Whinston, J. R. Green, et al. (1995). *Microeconomic theory*, Volume 1. Oxford university press New York.

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

Restuccia, D. and R. Rogerson (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic dynamics 11*(4), 707–720.

Rodríguez-Clare, A., M. Ulate, and J. P. Vasquez (2022). *Trade with nominal rigidities: Understanding the unemployment and welfare effects of the China shock*. CESifo, Center for Economic Studies & Ifo Institute.

Rubbo, E. (2020). Networks, phillips curves and monetary policy. Technical report, mimeo, Harvard University.

Samuelson, P. A. (1947). *Foundations of Economic Analysis*. Cambridge, MA: Harvard University Press.

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

Scitovszky, T. (1941). A note on welfare propositions in economics. *The Review of Economic Studies 9*(1), 77–88.

Solow, R. M. (1957). Technical change and the aggregate production function. *The review of Economics and Statistics*, 312–320.

Appendix

A Expenditure Function of Compensated Agent **45**

B Definition of Positive & Normative Representative Agent **46**

C Proofs **46**

D Additional Examples, Derivations, and Results **55**

E Explicit Characterization of Compensated Equilibrium **62**

E.1 Input-Output Notation 63

E.2 Differential Hat-Algebra 64

F Computation details: China shock example **67**

Appendix A Expenditure Function of Compensated Agent

The following proposition characterizes the expenditure function of the compensated agent.

Proposition 9 (Dual Representation of Compensated Representative Agent)**.** *The expenditure function associated with $U(\boldsymbol{c})$, in Theorem 1, denoted by $E(p, U)$ is*

$$E(\boldsymbol{p}, U) = \left(\sum_{h \in H} e_h(\boldsymbol{p}; u_h(\boldsymbol{c}_h^0)) \right) U.$$

By Shephard's lemma, the budget share of the compensated representative agent on good i, denoted b_i^{comp}, is

$$b_i^{comp}(\boldsymbol{p}) = \frac{\partial \log E(\boldsymbol{p}, U)}{\partial \log p_i} = \sum_h \frac{e_h(\boldsymbol{p}, u_h^0)}{\sum_{h'} e_{h'}(\boldsymbol{p}, u_{h'}^0)} b_{hi}(\boldsymbol{p}, u_h^0),$$

where $b_{hi}(\boldsymbol{p}, u_h^0)$ is the compensated budget share of household i at the status quo indifference curve $u_h^0 \equiv u_h(\boldsymbol{c}_h^0)$.

In words, the compensated agent's budget share on each good i is the average compensated budget share of all households, where each household is weighted according to its compensating income, $e_h(\boldsymbol{p}, u_h^0)$.

Given compensated aggregate budget shares, $b_i^{\text{comp}}(\boldsymbol{p})$, we can solve for equilibrium variables in the compensated equilibrium including prices $\boldsymbol{p}^{\text{comp}}$. Setting aggregate spending to be the numeraire in the compensated equilibrium, and using Theorem 1, we know that

$$A(t) = U(t) = \frac{1}{\sum_{h\in H} e_h(\boldsymbol{p}^{\text{comp}}(t); u_h(\boldsymbol{c}_h^0))} = A^{KH,comp}(t).$$

Appendix B Definition of Positive & Normative Representative Agent

We follow the definitions in Mas-Colell et al. (1995). We say that $u^{RA} : \mathbb{R}^N \to \mathbb{R}$ is a *positive representative agent* if the Marshallian demand curves generated by u^{RA}, given prices and total income, coincide with equilibrium allocations given the same prices and aggregate income:

$$\arg\max_{\boldsymbol{c}}\{u^{RA}(\boldsymbol{c}) : \sum_i p_i(t)c_i \le I(t)\} = \sum_h \arg\max_{\boldsymbol{c}_h}\{u_h(\boldsymbol{c}_h) : \sum_i p_i(t)c_{hi} \le I_h(t)\}.$$

The positive representative agent, $u^{RA} : \mathbb{R}^N \to \mathbb{R}$, is a *normative representative agent* relative to the social welfare function W if for every $(\boldsymbol{p}(t), I(t))$, the distribution of wealth across households, denoted by $\{I_h(t)\}$, also maximizes

$$W(v_1(\boldsymbol{p}(t), I_1(t)), \ldots, v_H(\boldsymbol{p}(t), I_H(t)))$$

subject to $\sum_{h=1}^H I_h(t) = I(t)$, where v_h is the indirect utility function of agent h.

Appendix C Proofs

Proof of Theorem 1. By (i), there exists an equilibrium allocation

$$\boldsymbol{c}^* \in \mathcal{C}\big(t, 1/A(t)\big) \quad \text{with} \quad \tilde{u}_h(\boldsymbol{c}_h^*) = 1 \text{ for all } h.$$

By the definition of $U(\boldsymbol{c}) = \min_h \tilde{u}_h(\boldsymbol{c}_h)$, this implies $U(\boldsymbol{c}^*) = 1$. Let $\boldsymbol{p}^*$ be an equilibrium price vector supporting $\boldsymbol{c}^*$ at $Z = 1/A(t)$, and define aggregate income

$$I^* \equiv \boldsymbol{p}^* \cdot \sum_h \boldsymbol{c}_h^*.$$

For each h, define the price index associated with the homothetic aggregator $\tilde{u}_h$ by

$$p_h^{CPI}(\boldsymbol{p}) \equiv \min_{\boldsymbol{c}_h}\{\boldsymbol{p}\cdot\boldsymbol{c}_h : \tilde{u}_h(\boldsymbol{c}_h) \geq 1\}.$$

We show that $\boldsymbol{c}_h^*$ is

$$\boldsymbol{c}_h^* \in \operatorname{argmin}_{\boldsymbol{c}_h}\{\boldsymbol{p}^*\cdot\boldsymbol{c}_h : \tilde{u}_h(\boldsymbol{c}) \geq 1\},$$

which implies that

$$\boldsymbol{p}^*\cdot\boldsymbol{c}_h^* = p_h^{CPI}(\boldsymbol{p}^*). \tag{18}$$

We prove this by contradiction. Suppose there exists some $\hat{\boldsymbol{c}}$ with $\tilde{u}_h(\hat{\boldsymbol{c}}_h) \geq 1$ and $\boldsymbol{p}^*\cdot\hat{\boldsymbol{c}}_h < \boldsymbol{p}^*\cdot\boldsymbol{c}_h^*$. By the definition of $\tilde{u}_h$, this implies

$$u_h(\hat{\boldsymbol{c}}_h) \geq u_h(\boldsymbol{c}_h(0)) = u_h(\boldsymbol{c}_h^*).$$

However, from utility maximization, we know that $\boldsymbol{c}_h^* \in \arg\max_{\boldsymbol{c}_h}\{u_h(\boldsymbol{c}_h) : \boldsymbol{p}^*\cdot\boldsymbol{c}_h \leq \boldsymbol{p}^*\cdot\boldsymbol{c}_h^*\}$. But if $u_h(\hat{\boldsymbol{c}}_h) \geq u_h(\boldsymbol{c}_h(0)) = u_h(\boldsymbol{c}_h^*)$ with $\boldsymbol{p}^*\cdot\hat{\boldsymbol{c}}_h < \boldsymbol{p}^*\cdot\boldsymbol{c}_h^*$, then by local nonsatiation, $\boldsymbol{c}_h^*$ could not be utility maximizing, which is a contradiction. This implies (18) holds, and summing it over h, gives

$$I^* = \sum_h \boldsymbol{p}^*\cdot\boldsymbol{c}_h^* = \sum_h p_h^{CPI}(\boldsymbol{p}^*). \tag{19}$$

Now, we show that the prices $\boldsymbol{p}^*$ and quantities $\boldsymbol{c}^*$ are also part of an equilibrium of the compensated economy with the as-if representative agent at $(t, Z = 1/A(t))$. Given a price vector $\boldsymbol{p}^*$, the compensated representative agent solves

$$\max_{\boldsymbol{c}} \min_h \tilde{u}_h(\boldsymbol{c}_h) \quad \text{s.t.} \quad \boldsymbol{p}^*\cdot(\sum_h \boldsymbol{c}_h) \leq I^*.$$

By two-stage budgeting, we can rewrite this problem in two steps. The as-if representative agent solves

$$\max_{\boldsymbol{\gamma}\geq 0} \min_h \gamma_h \quad \text{s.t.} \quad \sum_h \gamma_h p_h^{CPI}(\boldsymbol{p}^*) \leq I^*,$$

where each "good" γ_h has price $p_h^{CPI}(\boldsymbol{p}^*)$. Since $p_h^{CPI}(\boldsymbol{p}^*) \geq 0$ for all h,

$$\sum_h \gamma_h p_h^{CPI}(\boldsymbol{p}^*) \;\geq\; (\min_h \gamma_h) \sum_h p_h^{CPI}(\boldsymbol{p}^*),$$

so feasibility implies

$$\min_h \gamma_h \leq \frac{I^*}{\sum_h p_h^{CPI}(\boldsymbol{p}^*)} = 1,$$

where the final equality follows from (19). Hence $\min_h \gamma_h \leq 1$ for any feasible $\boldsymbol{\gamma}$. The choice $\gamma_h = 1$ for all h is affordable to the as-if representative agent since setting $\gamma_h^* = \tilde{u}_h(\boldsymbol{c}_h^*)$ yields:

$$\sum_h \gamma_h^* p_h^{CPI}(\boldsymbol{p}^*) = \sum_h p_h^{CPI}(\boldsymbol{p}^*) = I^*,$$

so the optimal value is $\min_h \gamma_h^* = 1$, and any optimal choice satisfies $\gamma_h^* = 1$ for all h. Therefore, $\boldsymbol{c}^*$ is an optimal choice for the as-if representative agent given prices $\boldsymbol{p}^*$ and income I^*. Since demand in the as-if representative agent economy coincides with demand in the compensated equilibrium, and all wedges and technologies are the same, the production quantities in the compensated equilibrium are also an equilibrium in the as-if representative agent economy. Hence, there is an equilibrium at $(t, Z = 1/A(t))$ with

$$\boldsymbol{c}^{\text{comp}}(t, 1/A(t)) = \boldsymbol{c}^* \quad \text{and} \quad U\big(\boldsymbol{c}^{\text{comp}}(t, 1/A(t))\big) = 1.$$

Finally, consider the compensated economy at $(t, Z = 1)$. In this economy there is a single representative agent with homothetic preferences U, constant-returns technologies, and wedges $\boldsymbol{\mu}(t)$ which, by (ii), are invariant to a change in factor-augmenting productivity Z. Hence,

$$\boldsymbol{c}^{\text{comp}}(t, 1/A(t)) = \boldsymbol{c}^{\text{comp}}(t, 1)/A(t).$$

Hence,

$$1 = U(\boldsymbol{c}^{\text{comp}}(t, 1/A(t))) = U(\boldsymbol{c}^{\text{comp}}(t, 1)/A(t)) = U(\boldsymbol{c}^{\text{comp}}(t, 1))/A(t).$$

where the last line follows from the fact that U is homogeneous of degree one. By definition $\boldsymbol{c}^{\text{comp}}(t) \equiv \boldsymbol{c}^{\text{comp}}(t, 1)$, so $A(t) = U\big(\boldsymbol{c}^{\text{comp}}(t)\big)$, which proves the first equality. Since the compensated representative agent has homothetic preferences, the remaining equalities follow as a consequence of standard results for representative agent economies.[31]

Consider now $t = 0$. By (i), the solution to $A(0)$ gives $U(\boldsymbol{c}^*) = 1$. Since the status-quo also gives $U(\boldsymbol{c}(0)) = 1$, $\boldsymbol{c}(0)$ is a solution to $A(0)$. Therefore, $A(0) = 1$ and $\boldsymbol{c}^{\text{comp}}(t, 1) = \boldsymbol{c}(0)$, which proves the last statement of the Theorem. □

Proof of Proposition 1. By Theorem 1, we know that

$$A(t) = Y^{comp}(t).$$

[31] The fact that Kaldor-Hicks efficiency, $A^{KH,comp}(t)$, in the equilibrium with a compensated representative agent coincides with the other measures follows trivially from the fact that this equilibrium has a single agent with homothetic preferences. It is important to note that $A^{KH,comp}(t)$ is not the same as $A^{KH}(t)$.

If preferences are identical, homothetic, and all households face the same relative prices, then the distribution of spending across households has no effect on equilibrium relative prices. Hence, the price and quantity of each good in the equilibrium with the compensated representative agent coincides with those in the decentralized equilibrium. That is, $\boldsymbol{p}^{comp}(t) = \boldsymbol{p}(t)$ and

$$\sum_h \boldsymbol{c}_h^{comp}(t) = \sum_h \boldsymbol{c}_h(t).$$

From this, it follows that

$$A(t) = Y^{comp}(t) = Y(t).$$

Since there is a positive representative agent with homothetic preferences, it follows from standard results (see, e.g., Baqaee and Burstein, 2023) that

$$Y(t) = A^{RA}(t).$$

Finally, letting $u(\boldsymbol{c})$ be the homogeneous of degree one representation of the utility function of every agent (since all agents have the same preferences), we have that

$$\begin{aligned} A^{KH}(t) &= \frac{\sum_h e_h(\boldsymbol{p}(t), u_h(t))}{\sum_h e_h(\boldsymbol{p}(t), u_h(0))} = \frac{\sum_h u_h(t)}{\sum_h u_h(0)} = \frac{\sum_h u(\boldsymbol{c}_h(t))}{\sum_h u(\boldsymbol{c}_h(0))}, \\ &= \frac{\sum_h u(\boldsymbol{c}(t)\chi_h(t))}{\sum_h u(\boldsymbol{c}(0)\chi_h(0))}, \end{aligned}$$

where $\chi_h(t)$ is household h's share of aggregate expenditures at t,

$$\begin{aligned} &= \frac{u(\boldsymbol{c}(t)) \sum_h \chi_h(t)}{u(\boldsymbol{c}(0) \sum_h \chi_h(0)}, \\ &= \frac{u(\boldsymbol{c}(t))}{u(\boldsymbol{c}(0))} = A^{RA}(t). \end{aligned}$$

□

Proof of Proposition 2. This follows from Hulten (1978). For any t, production functions are given by

$$y_i(t) = z_i(t) F_i(\{y_{ij}(t)\}_{j \in N}, \{l_{if}(t)\}_{f \in F}),$$

where F is the set of primary factors and N is the set of commodities. Resource constraints are

$$
\begin{aligned}
y_i(t) &= c_i(t) + \sum_{j \in N} y_{ji}(t) \\
z_f(t) L_f(t) &= \sum_{i \in N} l_{if}(t).
\end{aligned}
$$

By definition, the instantaneous change in real GDP for any $t > 0$ is given by:

$$
d \log Y = \sum_{i \in N} \frac{p_i(t) c_i(t)}{\sum_j p_j(t) c_j(t)} d \log c_i.
$$

Using the resource constraint

$$
d \log c_i = \frac{y_i}{c_i} d \log y_i - \sum_j \frac{y_{ji}}{c_i} d \log y_{ji}
$$

and the total derivative of the production function

$$
d \log y_i = d \log z_i + \sum_j \frac{\partial \log F_i}{\partial \log y_{ij}} d \log y_{ij} + \sum_f \frac{\partial \log F_i}{\partial \log l_{if}} d \log l_{if},
$$

we can write

$$
d \log c_i = \frac{y_i}{c_i} \left[d \log z_i + \sum_j \frac{\partial \log F_i}{\partial \log y_{ij}} d \log y_{ij} + \sum_f \frac{\partial \log F_i}{\partial \log l_{if}} d \log l_{if} \right] - \sum_j \frac{y_{ji}}{c_i} d \log y_{ji}
$$

Perfect competition implies that

$$
\frac{\partial \log F_i}{\partial \log y_{ij}} = \frac{p_j y_{ij}}{p_i y_i}.
$$

Hence,

$$
d \log c_i = \frac{y_i}{c_i} \left[d \log z_i + \sum_j \frac{p_j y_{ij}}{p_i y_i} d \log y_{ij} + \sum_f \frac{w_f l_{if}}{p_i y_i} d \log l_{if} \right] - \sum_j \frac{y_{ji}}{c_i} d \log y_{ji}
$$

Next, substitute this back into the definition of real GDP:

$$
\begin{aligned}
d\log Y &= \sum_{i\in N} \frac{p_i c_i}{\sum_j p_j c_j}\left[\frac{y_i}{c_i}\left[d\log z_i + \sum_j \frac{p_j y_{ij}}{p_i y_i} d\log y_{ij} + \sum_f \frac{w_f l_{if}}{p_i y_i} d\log l_{if}\right] - \sum_j \frac{y_{ji}}{c_i} d\log y_{ji}\right] \\
&= \sum_{i\in N} \frac{p_i y_i}{\sum_j p_j c_j} d\log z_i + \sum_{f\in F} \frac{w_f L_f}{\sum_j p_j c_j} d\log z_f,
\end{aligned}
$$

where the variables are all evaluated at t. The result is obtained by integrating. □

Proof of Proposition 3. This follows from combining Theorem 1 with Proposition 2. □

Proof of Corollary 1. The fact that $\Delta \log A \approx \sum_i \lambda_i(0)\Delta\log z_i$ is a consequence of Theorem 1 and Proposition 2. The fact that $\Delta\log A \approx \Delta\log Y$ is a consequence of Proposition 2. Finally, the fact that $\Delta\log Y \approx \Delta\log A^{KH}$ can be seen as follows.

$$
\begin{aligned}
\log A^{KH}(t) &= \log \frac{\sum_h e_h(\boldsymbol{p}(t), u_h(t))}{\sum_h e_h(\boldsymbol{p}(t), u_h(0))}, \\
&= \log \sum_h e_h(\boldsymbol{p}(t), u_h(t)) - \log \sum_h e_h(\boldsymbol{p}(t), u_h(0)), \\
&= \log \sum_i p_i(t)\left[\sum_h c_{hi}(t)\right] - \log \sum_h e_h(\boldsymbol{p}(t), u_h(0)), \\
d\log A|_{t=0} &= d\log \sum_i p_i(t)\left[\sum_h c_{hi}(t)\right] - d\log \sum_h e_h(\boldsymbol{p}(t), u_h(0)), \\
&= d\log \sum_i p_i(t)\left[\sum_h c_{hi}(t)\right] - \sum_i\sum_h \frac{e_h(\boldsymbol{p}(0), u_h(0))}{\sum_{h'} e_{h'}(\boldsymbol{p}(0), u_{h'}(0))}\frac{\partial \log e_h}{\partial \log p_i} d\log p_i, \\
&= \sum_i \frac{p_i(0)\left[\sum_h c_{hi}(0)\right]}{\sum_{h'} e_{h'}(\boldsymbol{p}(0), u_{h'}(0))} d\log p_i(t) + \sum_i \frac{p_i(0)\left[\sum_h c_{hi}(0)\right]}{\sum_{h'} e_{h'}(\boldsymbol{p}(0), u_{h'}(0))} d\log\left[\sum_{h'} c_{hi}(t)\right] \\
&- \sum_i \frac{p_i(0)\left[\sum_h c_{hi}(0)\right]}{\sum_{h'} e_{h'}(\boldsymbol{p}(0), u_{h'}(0))} d\log p_i, \\
&= \sum_i \frac{p_i(0)\left[\sum_h c_{hi}(0)\right]}{\sum_j p_j(0)\left[\sum_{h''} c_{hj}(0)\right]} d\log\left[\sum_{h'} c_{hi}(t)\right], \\
&= d\log Y,
\end{aligned}
$$

where we use the fact that $\sum_h e_h(\boldsymbol{p}(t), u_h(t)) = \sum_i p_i(t)\sum_h c_{hi}(t)$ for every $t \geq 0$ and we use Shephard's lemma to replace $\partial\log e_h/\partial\log p_i$ with $p_i(0)c_{hi}(0)/e_h(\boldsymbol{p}(0), u_h(0))$. Note that $\Delta\log Y \approx \Delta\log A^{KH}$ even if the initial equilibrium is distorted.

□

Proof of Corollary 2. This follows from combining Theorem 1 with Proposition 3 from Baqaee and Farhi (2019c). □

Proof of Proposition 4. Kaldor-Hicks efficiency $A^{KH}(t)$ can be defined analogously to $A(t)$ where we scale aggregate income instead of aggregate factor productivity. Specifically,

$$A^{KH}(t) = \max\left\{\phi \in \mathbb{R} : \text{there is } \boldsymbol{c} \in \phi^{-1}\mathcal{B}(\boldsymbol{p}(t), I(t)) \text{ and } u_h(\boldsymbol{c}_h) \geq u_h(\boldsymbol{c}_h^0) \text{ for every } h\right\},$$

where $\mathcal{B}(\boldsymbol{p}(t), I(t)) = \{\boldsymbol{c} : \boldsymbol{p}(t) \cdot \sum_h \boldsymbol{c_h} \leq I(t)\}$ and $I(t) = \sum_h I_h(t)$. In the absence of distortions ($\boldsymbol{\mu} = 1$), $\boldsymbol{p}(t)$ and $I(t)$ are prices and income in the competitive equilibrium.

We first show that $\mathcal{C}(\boldsymbol{z}(t), 1)$ is contained in the aggregate budget set $\mathcal{B}(\boldsymbol{p}(t), I(t))$. Suppose that there is a feasible consumption allocation $\boldsymbol{c}' \in \mathcal{C}(\boldsymbol{z}(t), 1)$ that violates the aggregate budget constraint at equilibrium prices. That is, $\boldsymbol{p}(t) \cdot \boldsymbol{c}' > I(t)$, where $I(t) = \boldsymbol{w}(t) \cdot \boldsymbol{L} + \Pi(t)$, and $\Pi(t)$ denotes aggregate profits which are equal to zero in equilibrium. Hence, $\Pi(t) < \boldsymbol{p}(t) \cdot \boldsymbol{c}' - \boldsymbol{w}(t) \cdot \boldsymbol{L}$. Aggregate profits Π' under any feasible allocation $\boldsymbol{c}'$ are given by $\Pi' = \boldsymbol{p}(t) \cdot \boldsymbol{c}' - \boldsymbol{w}(t) \cdot \boldsymbol{L}$. By the inequality above, $\Pi(t) < \Pi'$. This is a contradiction, since aggregate profits are maximized in a competitive equilibrium given prices (see Proposition 5.E.1 in Mas-Colell et al. (1995)).

By the second welfare theorem, the set $\mathcal{C}(\boldsymbol{z}(t), 1)$ is the Pareto frontier, which is contained in the set of all technologically feasible allocations, denoted $\mathcal{X}(t, 1)$. Furthermore, the feasible set $\mathcal{X}(t, 1)$ is contained in the aggregate budget set $\mathcal{B}(\boldsymbol{p}(t), I(t))$ as shown above. It follows that $A(t) \leq A^{KH}(t)$ because any choice $\boldsymbol{c}^*(t) \in \mathcal{C}(\boldsymbol{z}(t), 1/A(t)) \in \mathcal{X}(t, 1/A(t))$ that keeps every h at least indifferent to the status quo is also available by scaling $\mathcal{B}(\boldsymbol{p}(t), I(t)/A(t)$. Finally, pure redistributions leave the Pareto frontier unchanged, so $\mathcal{C}(t, 1)$ is unchanged with t and $A(t) = 1$. □

Proof of Proposition 5. This is a consequence of Theorem 1 and Petrin and Levinsohn (2012). Specifically, we can follow the derivation in Baqaee and Farhi (2019b). Suppressing the "comp" superscripts, we can write the following total differentials for any $t > 0$:

$$d \log c_i = \frac{y_i}{c_i} d \log y_i - \sum_{j \in N} \frac{y_{ji}}{c_i} d \log y_{ji},$$

and

$$\mu_j^{-1}(d \log y_j - d \log z_j - \sum_f \frac{w_f l_{jf}}{p_j y_j} \mu_j d \log l_{jf}) = \sum_{i \in N} \frac{p_i y_{ji}}{p_j y_j} d \log y_{ji}.$$

The first is an accounting identity and the second follows from cost-minimization. These two equations can be combined to obtain the desired result:

$$
\begin{aligned}
d\log Y &= \sum_{i\in N} \frac{p_i c_i}{\sum_j p_j c_j} d\log c_i, \\
&= \sum_{i\in N} \frac{p_i y_i}{\sum_j p_j c_j} d\log y_i - \sum_{i\in N}\sum_{j\in N} \frac{p_i y_{ji}}{\sum_j p_j c_j} d\log y_{ji}, \\
&= \sum_{i\in N} \frac{p_i y_i}{\sum_j p_j c_j} d\log y_i - \sum_{j\in N} \frac{p_j y_j}{\sum_j p_j c_j} \mu_j^{-1} (d\log y_j - d\log z_j - \sum_f \frac{w_f l_{jf}}{p_j y_j} \mu_j d\log l_{jf}), \\
&= \sum_{i\in N} \lambda_i \mu_i^{-1} d\log z_i + \sum_{i\in N} \lambda_i (1-\mu_i^{-1}) d\log y_i.
\end{aligned}
$$

Integrating this equation in the equilibrium with the compensated representative agent yields the desired result. □

Proof of Proposition 6. Index wedges by t, and denote the status quo with wedges by $\boldsymbol{c}^0(t)$. Let $t = 0$ denote the point where $\boldsymbol{\mu}(0) = \mathbf{1}$. The set $\mathcal{C}(0,1)$ is then the Pareto efficient frontier. From Proposition 5, we have that

$$
\log A(t) = -\int_0^t \sum_i \lambda_i^{\text{comp}}(s) \left(1 - \frac{1}{\mu_i^{\text{comp}}(s)}\right) \frac{d\log y_i^{\text{comp}}}{ds} ds.
$$

Differentiate the expression above to get

$$
\frac{d}{dt}[\log A] = -\sum_i \lambda_i^{\text{comp}}(t) \left(1 - \frac{1}{\mu_i^{\text{comp}}(t)}\right) \frac{d\log y_i^{\text{comp}}}{dt}.
$$

Differentiate a second time to get

$$
\begin{aligned}
\frac{d^2}{dt^2}[\log A] = -\sum_i d\lambda_i^{\text{comp}}(t) \left(1 - \frac{1}{\mu_i^{\text{comp}}(t)}\right) \frac{d\log y_i^{\text{comp}}}{dt} - \sum_i \lambda_i^{\text{comp}}(t) \frac{1}{\mu_i(t)} \frac{d\log \mu_i^{\text{comp}}}{dt} \frac{d\log y_i^{\text{comp}}}{dt} \\
- \sum_i \lambda_i^{\text{comp}}(t) \left(1 - \frac{1}{\mu_i^{\text{comp}}(t)}\right) \frac{d^2\log y_i^{\text{comp}}}{dt^2}.
\end{aligned}
$$

Evaluate these derivatives at $t = 0$ and write the second-order Taylor approximation:

$$
\log A \approx 0 - \frac{1}{2} \sum_i \lambda_i^{\text{comp}}(0) \frac{1}{\mu_i^{\text{comp}}(0)} \frac{d\log \mu_i^{\text{comp}}}{dt} dt \frac{d\log y_i^{\text{comp}}(0)}{dt} dt.
$$

To a second-order, this can also be written as

$$\Delta \log A \approx 0 - \frac{1}{2}\sum_i \lambda_i^{\text{comp}}(t)\frac{d\log \mu_i^{\text{comp}}}{dt}dt\frac{d\log y_i^{\text{comp}}(t)}{dt}dt \approx -\frac{1}{2}\sum_i \lambda_i^{\text{comp}}(t)\Delta \log \mu_i^{\text{comp}}\Delta \log y_i^{\text{comp}},$$

since the differences are higher-order. □

Proof of Proposition 8. Let $\boldsymbol{\tau}^*(t)$ and $\boldsymbol{T}^*(t)$ to be the maximizers of (16). Using $\boldsymbol{\tau}^*(t)$, we provide a slightly more general definition of the compensated equilibrium.

Definition 6 (Compensated Equilibrium)**.** An *equilibrium with a compensated representative agent* is the general equilibrium of an economy with the same technologies, resource constraints, wedges, and linear taxes $\boldsymbol{\tau}^*(t)$ as the original economy but where there is a representative agent with preferences $U(\boldsymbol{c})$.

The following lemma generalizes Theorem 1 to allow for limited redistribution.

Lemma 1 (Aggregate Efficiency Using Compensated Equilibrium)**.** *If condition* $(*)$ *holds and* $\boldsymbol{\mu}(t) = 1$*, then aggregate productivity can be calculated using the compensated equilibrium:*

$$A^{costly}(t) = Y^{comp}(t) = A^{KH,comp}(t) = A^{RA,comp}(t).$$

Moreover, at the status quo $t = 0$*, prices and quantities in the equilibrium with the compensated representative agent coincide with those in the decentralized equilibrium and* $A(0) = 1$*.*

Proof of lemma. Taking $\boldsymbol{\tau}^*(t)$ as given, we can apply Theorem 1 where $\boldsymbol{\tau}(t)$ are treated in the same way as the exogenous wedges $\boldsymbol{\mu}(t)$. Because condition $(*)$ holds, any constraints on transfers $\boldsymbol{T}^*(t)$ are non-binding, so the proof of Theorem 1 applies unchanged. □

In words, aggregate productivity with costly redistribution, $A^{\text{costly}}(t)$, can be computed by solving for changes in the welfare of the compensated representative agent (or equivalently, real GDP or Kaldor-Hicks efficiency in the equilibrium with such an agent). The only difference with Theorem 1 lies in knowing the necessary linear taxes $\boldsymbol{\tau}^*(t)$.

To conclude the proof of Proposition 8, combine Lemma 1 with Proposition 4 from Baqaee and Rubbo (2023) to obtain Equation (17).

□

Proof of Proposition 7. To obtain the first-order change in $\Delta \log A^{\text{costly}}(t)$, set the second order terms in Proposition 8 to zero. □

Proposition 10 (Equivalence of Kaldor-Hicks and $A(t)$)**.** *If there is one primary factor, so that relative prices are independent of demand, then* $A(t) = A^{KH}(t)$*.*

Proof. With one primary factor of production and constant-returns technologies, it is well-known that relative prices do not depend on final demand. Hence, the vector of equilibrium prices and aggregate income in the decentralized (multi-agent) economy $\boldsymbol{p}(t)$ and $I(t)$ are also equilibrium prices and aggregate income in the economy with a compensated representative agent. Theorem 1 implies that $A(t) = A^{KH,comp}(t)$. Since relative prices and aggregate income are the same, it follows that $A^{KH,comp}(t) = A^{KH}(t)$. □

Appendix D Additional Examples, Derivations, and Results

Individual consumption-equivalent for non-homothetic preferences. If $\succeq_h$ is non-homothetic, then $\tilde{u}_h$ is *not* a cardinalization of $\succeq_h$ (i.e. $\tilde{u}_h$ does not rank consumption allocations according to $\succeq_h$). Figure 3 graphically depicts indifference curves of $\tilde{u}_h$ — they are radial expansions of the status quo indifference curve defined by $u_h(\boldsymbol{c}_h) = u_h(\boldsymbol{c}_h(0))$. When $\succeq_h$ is homothetic, all indifference curves are radial expansions, so that the ranking produced by $\tilde{u}_h$ coincides with the one produced by $\succeq_h$.

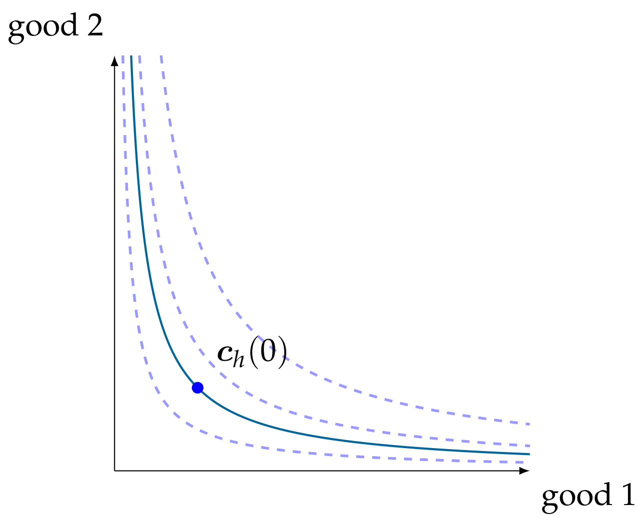


Figure 3: The solid blue line is the indifference curve $u_h(\boldsymbol{c}_h) = u_h(\boldsymbol{c}_h(0))$ and the dashed lines are the indifference curves of $\tilde{u}_h$.

Consider a household with non-homothetic CES preferences, as in Comin et al. (2021),

$$u_h(\boldsymbol{c}_h) = \left(\sum_i (c_{hi})^{\frac{\eta-1}{\eta}} (u_h(\boldsymbol{c}_h))^{\xi_i} \right)^{\frac{\eta}{\eta-1}}.$$

where η is the compensated elasticity of substitution and ξ_i controls income effects. Then

$\tilde{u}_h(\boldsymbol{c}_h)$ is homothetic CES given by

$$\tilde{u}_h(\boldsymbol{c}_h) = \frac{1}{u_h(0)} \left(\sum_i (c_{hi})^{\frac{\eta-1}{\eta}} (u_h(0))^{\xi_i} \right)^{\frac{\eta}{\eta-1}},$$

where $u_h(0) \equiv u_h(\boldsymbol{c}_h(0))$ is treated as a constant. If ξ_i are the same for every i, then $\tilde{u}_h$ and u_h are both cardinalizations of the same preference rankings.

Path-dependence of Real GDP Consider an economy with two Cobb-Douglas consumers and two goods. Each good is produced using a fixed, good-specific primary factor. Each household owns the same fraction of both factors, and denote this fraction by χ_h for household h. Let b_{hi} be consumer h's budget share on good i. Let $b_i = \chi_1 b_{1i} + \chi_2 b_{2i}$ be the aggregate budget share on good i.

Consider technology shocks to each good and shocks to the distribution of income between $t_0 = 0$ and $t_1 = 1$. Specifically, for $t \in (0, 1/2)$ set $\chi_1 = 2t$, $\chi_2 = 1 - 2t$, $(\log z_1, \log z_2) = (1 - (2t)^2, 1 - t)$. For $t \in (1/2, 1)$, set $\chi_1 = 2 - 2t$, $\chi_2 = 2t - 1$ and $(\log z_1, \log z_2) = (t, t)$. Hence, the economy starts and ends at the same point (allocations are identical at $t = 0$ and $t = 1$). For this reason, $A(t) = 1$.

However, real GDP $Y(t) \neq 1$. To see this:

$$\Delta \log Y = \int_{t_0}^{t_1} b_1 d\log c_1 + b_2 d\log c_2 = \int_{t_0}^{t_1} \left(b_1 \frac{d\log z_1}{dt} + b_2 \frac{d\log z_2}{dt} \right) dt = \frac{b_{21} - b_{11}}{12}.$$

Hence, as long as $b_{11} \neq b_{21}$, i.e. preferences are not the same, real GDP is non-zero. It can be increased to be arbitrarily high by running the loop in one direction multiple times and arbitrarily low by running the loop in reverse. Of course, this means that if we add a small positive productivity shock after running these loops, then $Y(t)$ can be lower than $Y(0)$ even though there is more of every consumption good and every household is strictly better off.

Derivations in Example 2 The individual consumption-equivalent function associated to

$$u_h(\boldsymbol{c}_h) = c_{hg}^{\alpha} c_{hs}^{1-\alpha},$$

is

$$\tilde{u}_h(\boldsymbol{c}_h) = \frac{c_{hg}^{\alpha} c_{hs}^{1-\alpha}}{\left(c_{hg}^{0}\right)^{\alpha} \left(c_{hs}^{0}\right)^{1-\alpha}}.$$

The compensated representative agent assuming interior outcomes sets

$$A = \tilde{u}_h = \tilde{u}_{h'}$$

for all h' subject to

$$\sum_h c_{hg} = z_g, \qquad c_{hs} = z_{hs}.$$

Hence, $\{c_{hg}\}_{h=1}^{H}$ solves

$$A = \frac{c_{hg}^{\alpha} z_{hs}^{1-\alpha}}{\left(c_{hg}^{0}\right)^{\alpha}\left(c_{hs}^{0}\right)^{1-\alpha}} = \frac{c_{h'g}^{\alpha} z_{h's}^{1-\alpha}}{\left(c_{h'g}^{0}\right)^{\alpha}\left(c_{h's}^{0}\right)^{1-\alpha}}, \qquad \text{for all } h'$$

subject to $\sum_h c_{hg} = z_g$. The solution is

$$A = \left(\frac{z_g}{\sum_h z_{hs}^{\frac{\alpha-1}{\alpha}} c_{hg}^{0} \left(z_{hs}^{0}\right)^{\frac{1-\alpha}{\alpha}}}\right)^{\alpha}.$$

In status quo, $c_{hg}^{0} = \chi_h^0 z_g^0$,[32] so

$$A = \left(\frac{z_g / z_g^0}{\sum_h \chi_h^0 \left(z_{hs}/z_{hs}^0\right)^{\frac{\alpha-1}{\alpha}}}\right)^{\alpha}.$$

A second-order approximation yields the expression in the text.

Derivations in Example 3 The individual consumption-equivalent function associated to the utility function

$$u_h(\boldsymbol{c}_h) = \left[(\alpha_h)^{\frac{1}{\theta_h}} \left(u_h(\boldsymbol{c}_h)\right)^{\zeta_h} c_{hd}^{\frac{\theta_h-1}{\theta_h}} + (1-\alpha_h)^{\frac{1}{\theta_h}} c_{hf}^{\frac{\theta_h-1}{\theta_h}}\right]^{\frac{\theta_h}{\theta_h-1}},$$

is

$$\tilde{u}_h(\boldsymbol{c}_h) = \frac{1}{u_h^0}\left[(\alpha_h)^{\frac{1}{\theta_h}} \left(u_h^0\right)^{\zeta_h} c_{hd}^{\frac{\theta_h-1}{\theta_h}} + (1-\alpha_h)^{\frac{1}{\theta_h}} c_{hf}^{\frac{\theta_h-1}{\theta_h}}\right]^{\frac{\theta_h}{\theta_h-1}},$$

[32]To see that χ_h^0 is also region h's share in total income, note that (under the Cobb-Douglas specification of this example) the first-order condition for c_{hg} and c_{hs} is $p_{hs}^0 z_{hs}^0 = \frac{1-\alpha}{\alpha}\chi_h^0 z_g^0$, where we normalize the price of the tradable good to 1. Therefore, $\chi_h^0 z_g^0 + p_{hs}^0 z_{hs}^0 = \frac{1}{\alpha}\chi_h^0 z_g^0$ and $\frac{\chi_h^0 z_g^0 + p_{hs}^0 z_{hs}^0}{\sum_{h'} \chi_{h'}^0 z_g^0 + p_{h's}^0 z_{h's}^0} = \frac{\chi_h^0}{\sum \chi_{h'}^0} = \chi_h^0$.

In autarky, $c_{hf} = 0$, so

$$\tilde{u}_h([c_{hd}, 0]) = \frac{1}{u_h^0} c_{hd} \left[(\alpha_h)^{\frac{1}{\theta_h}} (u_h^0)^{\zeta_h} \right]^{\frac{\theta_h}{\theta_h - 1}}$$

The domestic expenditure share of household h in the status quo is

$$s_{hd}^0 \equiv \frac{p_d^0 c_{hd}^0}{p_d^0 c_{hd}^0 + p_f^0 c_{fd}^0} = \frac{p_d^0 c_{hd}^0}{p_h^0 c_h^0} = \frac{(\alpha_h)^{\frac{1}{\theta_h}} (u_h^0)^{\zeta_h} \left(c_{hd}^0\right)^{\frac{\theta_h - 1}{\theta_h}}}{(u_h^0)^{\frac{\theta_h - 1}{\theta_h}}}$$

Solving for $\left[(\alpha_h)^{\frac{1}{\theta_h}} (u_h^0)^{\zeta_h} \right]$ and substituting into the individual consumption-equivalent function yields

$$\begin{aligned}
\tilde{u}_h([c_{hd}, 0]) &= \frac{c_{hd}}{c_{hd}^0} \left(s_{hd}^0\right)^{\frac{\theta_h}{\theta_h - 1}} \\
&= \frac{p_d^0}{p_h^0 c_h^0} \frac{p_h^0 c_h^0}{p_d^0 c_{hd}^0} c_{hd} \left(s_{hd}^0\right)^{\frac{\theta_h}{\theta_h - 1}} \\
&= \frac{p_d^0}{p_h^0 c_h^0} c_{hd} \left(s_{hd}^0\right)^{\frac{1}{\theta_h - 1}} \\
&= \frac{p_d^0 y_d^0}{p_h^0 c_h^0} \frac{c_{hd}}{y_d^0} \left(s_{hd}^0\right)^{\frac{1}{\theta_h - 1}} \\
&= \frac{1}{\chi_h^0} \frac{c_{hd}}{y_d^0} \left(s_{hd}^0\right)^{\frac{1}{\theta_h - 1}},
\end{aligned}$$

where $\chi_h^0 = p_h^0 c_h^0 / p_d^0 y_d^0$ is the share of $h's$ expenditures in total income (assuming balanced trade), and y_d^0 is the aggregate quantity of the home produced good in the status quo (which is consumed and exported).

The compensated representative agent assuming interior outcomes sets

$$\tilde{u}_h = \tilde{u}_{h'}$$

for all h' subject to

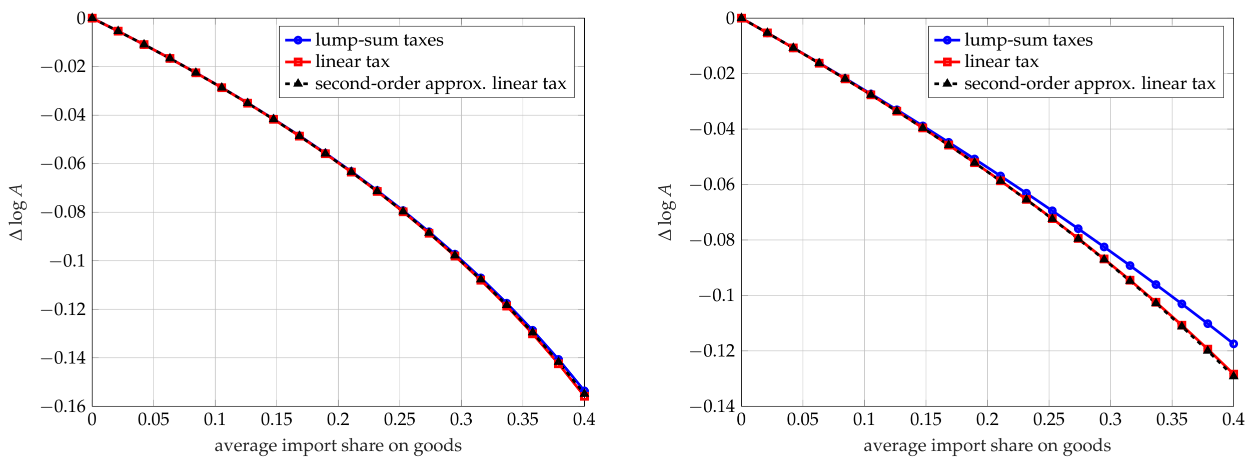
$$\sum c_{hd} = y_d = y_d^0,$$

where y_d is the total output of domestic good in autarky (which is equal to that in status quo). Combining, we obtain

$$\Delta \log A = \log \tilde{u}_h = \log \frac{1}{\sum \chi_h^0 \left(s_{hd}^0\right)^{-\frac{1}{\theta_h - 1}}},$$

which is the expression in the text.

Numerical version of Example 8 Figure 4 numerically illustrates the performance of the second-order approximation to the exact solution with distortionary linear taxes, and compares them both to the solution with lump-sum taxes. The second-order approximation performs well even for large shocks. Panel (a) uses $\rho = 0.5$, so consumption and leisure are complements and the Frisch elasticity of labor supply is a reasonable 0.5. Since ρ is low, distortionary taxes are able to achieve an outcome that is roughly as good as lump-sum taxes. Panel (b) uses a much higher $\rho = 3$. In this case, the gap between the lump-sum and linear taxation scenarios is larger since consumption taxes reduces labor and increase leisure, which causes efficiency to fall.



(a) $\rho = 0.5$ (b) $\rho = 3.0$

Figure 4: A numerical example of the losses from autarky with and without distortionary redistribution. The other parameter values are $\Omega_d = 0.5$, $\chi_h = 0.5$, $\theta_h = 3$, $s_{hd} = 3s_{h'd}$.

Example with skill-biased technical change and costly redistribution. We now consider a simple example with skill-biased technical change that raises the real wage of high-skill workers but lowers the real-wage for low-skill workers. We compare how the response of aggregate efficiency changes depending on the redistributive tools available.

Suppose that output (and consumption) are a CES aggregate of the output of manufacturing and services:

$$c = y = \left[\gamma_1^{\frac{1}{\rho}} y_m^{\frac{\rho-1}{\rho}} + (1-\gamma_1)^{\frac{1}{\rho}} y_s^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}},$$

where each sector's output is a CES aggregate of low- and high-skill labor

$$y_o = \left[\alpha_o^{\frac{1}{\sigma}} \left(z_{o1} l_{o1} \right)^{\frac{\sigma-1}{\sigma}} + (1-\alpha_o)^{\frac{1}{\sigma}} \left(z_{o2} l_{o2} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}},$$

where l_{o1} is low- and l_{o2} is high-skill labor. The resource constraints are that

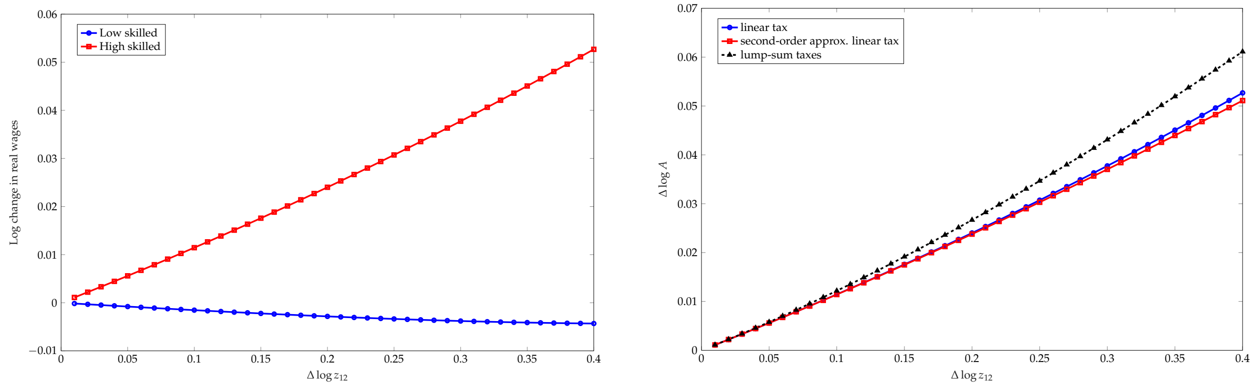
$$\sum_h c_h = c, \qquad \sum_{o=\{m,s\}} l_{o1} = l_1, \qquad \sum_{o=\{m,s\}} l_{o2} = l_2.$$

We assume that workers are much more substitutable than sectors: $\rho \ll \sigma$. We also assume that manufacturing is more intensive in low-skill labor use than services.

Consider an increase in automation or the productivity of capital, which we capture via an increase in the productivity of high-skill labor in manufacturing: $\Delta \log z_{m2} > 0$. This is a reduced-form representation for the idea that high-skill labor in manufacturing is equipped by capital, and hence an increase in the quality of capital makes high-skill more productive.[33]

Again, we contrast two scenarios: (1) lump-sum taxation is available, (2) lump-sum transfers must be non-negative and the government can only levy a linear tax on machine use in manufacturing, which we capture as a linear tax, τ, on manufacturing's use of high-skill labor. Figure 5 illustrates the results in a numerical example. Panel 5a shows that skill-biased technical change raises the real wage for high-skill workers and lowers them for low-skill workers in the decentralized equilibrium. The fact that low-skill wages decline means that they need to be compensated via transfers financed by either lump-sum or distortionary taxes. Panel 5b shows the increase in efficiency depending on which taxes are used. As expected, the increase in aggregate efficiency is lower if only distortionary redstributive tools are available. Panel 5b also shows that the second-order approximation is very accurate. In the absence of any redistributive tools whatsoever, aggregate efficiency in this example actually declines because the low-skill workers are worst off and there is no feasible way to compensate them.

[33]For example, high-skill labor and capital are combined in a Leontief nest together called equipped labor, and then equipped labor is substitutable with low-skill labor. We can then think of altering the productivity of equipped labor by varying the productivity of capital.



(a) real wages (b) aggregate efficiency

Figure 5: A numerical example of skill-biased technical change. The parameter values are $\rho = 1$, $\sigma = 8$, $\gamma = 0.5$, $\alpha_{m1} = 0.9$, and $\alpha_{s1} = 0.5$. We normalize steady-state quantities so that the CES share parameters are equal to expenditure shares in the status quo.

Aggregate productivity defined using technologically feasible allocations Define the set of technologically feasible allocations, $\mathcal{C}^F(t,Z)$, as the set of consumption allocations that are compatible with production technologies and resource constraints at (t,Z):

$$\mathcal{C}^F(t,Z) \equiv \left\{ \boldsymbol{c} \in \mathbb{R}^{H\times N} : \begin{array}{l} \text{there exist } (y_i, y_{ij}, l_{if}) \text{ such that} \\ y_i = z_i(t)\, F_i(\{y_{ij}\}_j,\ Z\{l_{if}\}_f) \text{ for all } i, \\ \sum_j y_{ji} + \sum_h c_{hi} \le y_i \text{ for all } i, \\ \sum_i l_{if} \le z_f(t)\, L_f \text{ for all } f \end{array} \right\}.$$

Using this set, we can define a measure of aggregate productivity that expands or contracts $\mathcal{C}^F(t,Z)$:[34]

$$A^F(t) \equiv \max \left\{ Z \in \mathbb{R} : \text{there is } \boldsymbol{c} \in \mathcal{C}^F(t, 1/Z) \text{ with } \boldsymbol{c}_h \succeq_h \boldsymbol{c}_h(0) \text{ for every } h \right\}.$$

The following proposition shows that, in the absence of distortions, $A^F(t)$ coincides with the measure $A(t)$ defined over allocations that can be supported as equilibria with lump-sum transfers.

Proposition 11 (Equivalence of A using Feasible and Equilibrium Allocations). *If* $\boldsymbol{\mu}(t) \equiv$

[34]Because each F_i has constant returns to scale and the resource constraints are linear, the technologically feasible set scales linearly with the factor-augmenting shifter: for any $Z > 0$, $\mathcal{C}^F(t,Z) = Z\,\mathcal{C}^F(t,1)$. Hence scaling Z in the definition of $A^F(t)$ is equivalent to radially scaling the feasible set $\mathcal{C}^F(t,1)$ in consumption space.

$\mathbf{1}$ *for* $t > 0$*, then* $A(t) = A^F(t)$*.*

Proof. First, any decentralized equilibrium allocation is technologically feasible, so $\mathcal{C}(t, Z) \subseteq \mathcal{C}^F(t, Z)$ for all Z. Hence any Z feasible in the definition of $A(t)$ is also feasible for $A^F(t)$, and therefore $A^F(t) \geq A(t)$.

For the reverse inequality, let $Z = A^F(t)$ and let $\boldsymbol{c} \in \mathcal{C}^F(t, 1/Z)$ be an allocation that attains $A^F(t)$, i.e. $\boldsymbol{c}_h \succeq_h \boldsymbol{c}_h(0)$ for all h. If $\boldsymbol{c}$ were not Pareto efficient in $\mathcal{C}^F(t, 1/Z)$, there would exist $\boldsymbol{c}' \in \mathcal{C}^F(t, 1/Z)$ that Pareto dominates $\boldsymbol{c}$, and by continuity and local nonsatiation we could then slightly increase the contraction factor above Z while still keeping everyone at least as well off as at $\boldsymbol{c}(0)$, contradicting the maximality of $Z = A^F(t)$. Thus $\boldsymbol{c}$ is Pareto efficient in $\mathcal{C}^F(t, 1/Z)$ and $\boldsymbol{c}_h \succeq_h \boldsymbol{c}_h(0)$ for all h.

Since there are no wedges and transfers are lump sum, the second welfare theorem implies that any Pareto-efficient allocation in $\mathcal{C}^F(t, 1/Z)$ can be decentralized as a Walrasian equilibrium with some prices and transfers. In particular, $\boldsymbol{c}$ can be decentralized in this way, so $\boldsymbol{c} \in \mathcal{C}(t, 1/Z)$ and $\boldsymbol{c}_h \succeq_h \boldsymbol{c}_h(0)$ for all h. Therefore Z is also feasible in the definition of $A(t)$, implying $A(t) \geq A^F(t)$.

Combining $A^F(t) \geq A(t)$ and $A(t) \geq A^F(t)$ yields $A(t) = A^F(t)$. □

Appendix E Explicit Characterization of Compensated Equilibrium

Theorem 1 and Lemma 1 (in Appendix C) show that calculating changes in aggregate productivity can be boiled down to solving for the equilibrium of an economy with a compensated representative agent. This section provides some formulas for calculating variables in the compensated equilibrium. To do so, we rely on the differential hat algebra approach in Baqaee and Farhi (2020), which characterizes equilibria of representative agent economies with wedges using differential equations. Alternatively, one could also use exact-hat algebra methods, as in Dekle et al. (2008).

For concreteness, assume that all production and utility functions are nested-CES. (Non-CES economies can be analyzed in a similar way following the non-CES extensions in Baqaee and Farhi (2019c)). To make the notation more compact, represent the economy in such a way that each producer, i, is associated with a single elasticity of substitution θ_i (by treating each sub-nest as a separate producer).

E.1 Input-Output Notation

Stack the expenditure shares of the representative household, all producers, and all factors into the $(H+N+F) \times (H+N+F)$ input-output matrix Ω. The first H rows correspond to the households consumption baskets. The next N rows correspond to the expenditure of each producer on every other producer and factor as a share of its sales (where the sales price is always inclusive of the wedge and tax). The last F rows correspond to the expenditure shares of the primary factors (which are all zeros, since primary factors do not require any inputs). With some abuse of notation, the heterogeneous agent input-output matrix can be written as

$$\Omega = \left[\begin{array}{ccc|ccc|ccc} 0 & \cdots & 0 & b_{11} & \cdots & b_{1N} & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & & \cdots & & & \cdots & \\ 0 & \cdots & 0 & b_{H1} & \cdots & b_{HN} & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & \Omega_{11} & \cdots & \Omega_{1N} & \Omega_{1N+1} & \cdots & \Omega_{1N+F} \\ \vdots & \cdots & \vdots & & \ddots & & & & \\ 0 & \cdots & 0 & \Omega_{N1} & & \Omega_{NN} & \Omega_{NN+1} & \cdots & \Omega_{NN+F} \\ \hline 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right].$$

Note that our convention is that rows (not columns) record costs relative to revenues inclusive of wedges and taxes. If wedges and taxes are greater than one, then the rows of this matrix will generally sum to a number less than one. The Leontief inverse matrix is the $(H+N+F) \times (H+N+F)$ matrix defined as

$$\Psi \equiv (I-\Omega)^{-1} = I + \Omega + \Omega^2 + \ldots,$$

where I is the identity matrix. The Leontief inverse matrix $\Psi \geq I$ records the *direct and indirect* exposures through the supply chains in the production network.

Denote the distribution of expenditures by each household by χ, which is an $(H+N+F) \times 1$ vector. The first H elements are equal to each household's share of aggregate consumption expenditures, and the remaining $N+F$ elements are all zeros. As a matter of accounting identities, the vector of Domar weights satisfies:

$$\lambda' = \chi' \Psi.$$

In this equation λ is a $(H+N+F)\times 1$ vector. The first H elements are equal the expenditures of each household relative to aggregate consumption expenditures, χ', the next $N+F$ elements are equal to the sales of each good and factor relative to aggregate consumption expenditures.

Let μ and τ denote the diagonal matrices whose iith element is equal to μ_i and τ_i respectively. Recall that μ are exogenous wedges, whereas τ are linear taxes that can be used for redistribution. Define the cost-based Leontief inverse to be

$$\tilde{\Psi} = (I - (\tau\mu)\Omega)^{-1}.$$

Note that the cost-based Leontief inverse coincides with Ψ in the absence of wedges. Intuitively, $\tilde{\Psi}$ is a version of the Leontief inverse that calculates exposures of i to j in terms of cost shares rather than revenue shares (revenues exceed costs if wedges and taxes are greater than one).

For any non-negative vector a, define

$$Cov_a(b,c) = \mathbb{E}_a[bc] - \mathbb{E}_a[b]\mathbb{E}_a[c] = \sum_i \frac{a_i}{\sum_{i'} a_{i'}} b_i c_i - \sum_i \frac{a_i}{\sum_{i'} a_{i'}} b_i \sum_i \frac{a_i}{\sum_{i'} a_{i'}} c_i,$$

where $\mathbb{E}_a[\cdot]$ denotes averages of vectors weighted by the elements of a. For any matrix X, denote its ith row and column by $X_{(i,:)}$ and $X_{(:,i)}$.

E.2 Differential Hat-Algebra

The next proposition characterizes compensated variables in terms of initial expenditure shares, wedges, and shocks.

Proposition 12 (Differential Equations for Compensated Equlibrium)**.** *Let aggregate spending be the numeraire. Then, assuming the conditions of Theorem 1 or Lemma 1 hold, the compensated equilibrium satisfies the following system of differential equations. For each $i \in H+N+F$, the compensated price satisfies*

$$d\log p_i^{comp} = \sum_j \tilde{\Psi}_{ij}^{comp}[d\log \mu_j \tau_j^* - d\log z_j] + \sum_{f\in F} \tilde{\Psi}_{if}^{comp} d\log \lambda_f^{comp}. \tag{20}$$

Compensated Domar weights for goods and factors satisfy

$$d\lambda_l^{comp} = \sum_j \lambda_j^{comp}(1-\theta_j)\mu_j^{-1} Cov_{\Omega_{(j,:)}^{comp}}\left(d\log p^{comp}, \Psi_{(:,l)}^{comp}\right) + Cov_{\chi^{comp}}\left(d\log \chi^{comp}, \Psi_{(:,l)}^{comp}\right) - \sum_j \lambda_j^{comp}\left(\Psi_{jl} - \mathbf{1}[j=l]\right) d\log \mu_j \tau_j^*. \quad (21)$$

Changes in compensated expenditure shares for household h satisfy

$$d\log \chi_h^{comp} = d\log p_h^{comp}, \quad (22)$$

where $d\log p_h^{comp}$ is the price of the consumption bundle for household h. The compensated input-output matrix satisfies

$$d\Omega_{ij}^{comp} = (1-\theta_i)\left(d\log p_j^{comp} - \mathbb{E}_{\Omega_{(i,:)}^{comp}}[d\log p^{comp}]\right) - d\log \mu_i. \quad (23)$$

Finally, $d\log y_i^{comp}$ is given by $d\log \lambda_i^{comp} - d\log p_i^{comp}$. The initial conditions are that all prices and expenditures are equal to the ones in the status quo decentralized equilibrium for $t = 0$.

Equation (20), (21), and (23) are standard and identical to expressions in Baqaee and Farhi (2020). They are loglinearizations of marginal cost-functions, market clearing conditions, and demand curves respectively. The key equation, which distinguishes the compensated equilibrium from the decentralized equilibrium is (22). Whereas in the decentralized equilibrium changes in household expenditures are determined by changes in the income of each household, in the compensated equilibrium, they are determined by the choices of the compensated agent (who tries to equate individual consumption-equivalents across agents). The term $d\log p_h^{\text{comp}}$, which is pinned down by (20), is the change in the compensated price index of household h.

The taxes $\boldsymbol{\tau}^*(t)$ are given by the maximizers of the problem in (16). If only lump-sum transfers are used for redistribution, as in Sections 4 and 5, then $\boldsymbol{\tau}^*(t) = 0$, and Proposition 12 fully characterizes the compensated equilibrium in terms of exogenous parameters: $\boldsymbol{z}(t)$ and $\boldsymbol{\mu}(T)$. If lump-sum transfers are unavailable, then solving for $\tau^*(t)$ requires specifying more details about the set of available tax instruments. Specifically, we would need to add the log-linearized first-order conditions for the tax instruments from (16) as additional equations in Proposition 12 to pin down how τ^* evolves.

There is one case where this optimization problem can be avoided. If there are only $H-1$ taxes available, then (22) can pin down $\tau^*(t)$. For example, suppose that there are $H-1$ taxes, and the share of revenues from the ith tax sent to household h are given by

α_{ih}:

$$T_h(t) = \sum_i \alpha_{ih} \left(1 - \frac{1}{\tau_i^*(t)}\right) \lambda_i(t).$$

Log-differentiating household h's budget constraint gives:

$$d \log \chi_h^{\text{comp}} = \sum_f \frac{\omega_{hf} \lambda_f^{\text{comp}}}{\chi_h^{\text{comp}}} d \log \lambda_f^{\text{comp}} + \frac{dT_h}{\chi_h^{\text{comp}}},$$

differentiating $T_h(t)$ above, and substituting it into the log-linearized budget constraint gives $H - 1$ additional equations which, assuming regularity conditions, will pin down $d \log \boldsymbol{\tau}^*$.

Generally, solving the system of linear equations in Proposition 12 requires inverting a system of equations. When there is a single primary factor of production and we evaluate these derivatives at a perfectly competitive point, then the change in efficiency can be solved out easily up to a second-order.

Proposition 13 (Aggregate Efficiency with One Factor)**.** *Consider a competitive economy with a single primary factor of production. The change in aggregate efficiency in response to a vector of productivity shocks,* $\Delta \log \boldsymbol{z}$ *and changes in wedges* $\Delta \log \boldsymbol{\mu}$ *is*

$$\Delta \log A \approx \sum_i \lambda_i \Delta \log z_i + \frac{1}{2} \sum_{i \in N+H} \lambda_i (\theta_i - 1) Var_{\Omega(i,:)} \left(\sum_k \Psi_{(:,k)} \log z_k \right)$$
$$- \frac{1}{2} \sum_{i \in N+H} \lambda_i \theta_i Var_{\Omega(i,:)} \left(\sum_k \Psi_{(:,k)} \Delta \log(\mu_k \tau_k^*) \right).$$

to a second-order approximation in $\Delta \log z$ *and* $\Delta \log \mu$.

There are three summands. The first one is just Hulten's theorem. The second summand is a nonlinear adjustment due to changes in Domar weights. The second summand is also equal to: $1/2 \sum_k \left[\sum_j \partial \lambda_k^{\text{comp}} / \partial \log z_j \Delta \log z_j \right] \Delta \log z_k$. If the compensated Domar weight for k rises due to productivity shocks, then the shock to k is more important. This happens if exposure to k is heterogeneous, captured by the variance term, and if elasticities of substitution, θ_i, are far from unity. The final summand are the Harbeger triangles caused by the taxes and wedges. The triangles are larger the higher are elasticities of substitution, θ_i, and the more heterogeneous are exposures to the taxes and wedges, captured by the variance terms.

Appendix F Computation details: China shock example

Before providing details for our China shock example, we discuss more generally how to compute $A^{\text{costly}}(t)$ when some agents are not made exactly indifferent to status-quo (in which case Theorem 1 does not apply). The following proposition expresses the problem of solving $A^{\text{costly}}(t)$ as a maximization problem for the case when preferences are homothetic.

Proposition 14 (Calculating A with Costly Redistribution and Homothetic Preferences)**.** *Suppose that preferences are homothetic for all h, and wedges $\boldsymbol{\mu}(t)$ are invariant to Z in equilibrium. Then*

$$A^{costly}(t) = \max_{\boldsymbol{c} \in \mathcal{C}^{costly}(t,1)} \min_h \tilde{u}_h(\boldsymbol{c}_h). \tag{24}$$

Proof. We first note that

$$\mathcal{C}^{\text{costly}}(t, Z) = Z\, \mathcal{C}^{\text{costly}}(t, 1). \tag{25}$$

To see this, consider a consumption allocation $\boldsymbol{c} \in \mathcal{C}^{\text{costly}}(t, Z)$, which is an equilibrium supported by wedges $\boldsymbol{\mu}$, taxes $\boldsymbol{\tau}$, transfers (relative to total spending) $\boldsymbol{T}$, and prices $\boldsymbol{p}$. Using the two assumptions in the proposition and constant returns to scale in production, one can verify that, with factor-augmenting technology Z', the consumption allocation $\frac{Z'}{Z}\boldsymbol{c} \in \mathcal{C}^{\text{costly}}(t, Z')$ is an equilibrium supported with the same wedges $\boldsymbol{\mu}$, taxes $\boldsymbol{\tau}$, transfers (relative to total spending) $\boldsymbol{T}$, and prices $\boldsymbol{p}$. Setting $Z' = 1$, we obtain that if $\boldsymbol{c} \in \mathcal{C}^{\text{costly}}(t, Z)$, then $\frac{1}{Z}\boldsymbol{c} \in \mathcal{C}^{\text{costly}}(t, 1)$, i.e. $\boldsymbol{c} \in Z\mathcal{C}^{\text{costly}}(t, 1)$. Reversing the argument shows the converse inclusion, so Equation (25) follows.

Define the value of $\mathcal{C}^{\text{costly}}(t, Z)$ for the compensated representative agent by

$$V(t, Z) = \max_{\boldsymbol{c} \in \mathcal{C}^{\text{costly}}(t,Z)} U(\boldsymbol{c}), \quad \text{where} \quad U(\boldsymbol{c}) = \min_h \tilde{u}_h(\boldsymbol{c}_h).$$

By Equation (25) and homogeneity of degree 1 of $U(\boldsymbol{c})$, we have

$$V(t, Z) = Z\, V(t, 1). \tag{26}$$

By definition of $A^{\text{costly}}(t)$, there exists

$$\boldsymbol{c}^* \in \mathcal{C}^{\text{costly}}\left(t, \frac{1}{A^{\text{costly}}(t)}\right)$$

with $\tilde{u}_h(\boldsymbol{c}_h^*) \geq 1$ for all h (equivalently, $\boldsymbol{c}_h^* \succeq_h \boldsymbol{c}_h(0)$ for all h). Hence $U(\boldsymbol{c}^*) = \min_h \tilde{u}_h(\boldsymbol{c}_h^*) \geq$

1, and therefore $V\big(t, 1/A^{\text{costly}}(t)\big) \geq 1$. Using (26), this implies

$$V(t,1) \geq A^{\text{costly}}(t).$$

Conversely, suppose $V(t,1) > A^{\text{costly}}(t)$. Choose a contraction factor X such that $A^{\text{costly}}(t) < X < V(t,1)$. Then, by (26),

$$V\Big(t, \frac{1}{X}\Big) = \frac{1}{X} V(t,1) > \frac{1}{X} X = 1.$$

Hence there exists $\boldsymbol{c} \in \mathcal{C}^{\text{costly}}(t, 1/X)$ with $U(\boldsymbol{c}) \geq 1$, which implies $\tilde{u}_h(\boldsymbol{c}_h) \geq 1$. By the definition of $A^{\text{costly}}(t)$, such an X is then a feasible contraction factor, contradicting the maximality of $A^{\text{costly}}(t)$. Thus $V(t,1) \leq A^{\text{costly}}(t)$. Combining the two inequalities, we conclude that $V(t,1) = A^{\text{costly}}(t)$, which together with the definition of $V(t,1)$ and U yields the desired expression for $A^{\text{costly}}(t)$. □

We can make the max–min problem in Equation (24) more computationally tractable by rewriting it as

$$A^{\text{costly}}(t) = \max\left\{x : (\boldsymbol{\tau}, \boldsymbol{T}) \in \mathcal{T},\ \tilde{u}_h\big(\boldsymbol{c}_h(t, \boldsymbol{\tau}, \boldsymbol{T})\big) \geq x \text{ for all } h\right\}, \tag{27}$$

where $\boldsymbol{c}_h(t, \boldsymbol{\tau}, \boldsymbol{T})$ denotes agent h's consumption in the equilibrium with productivities $\boldsymbol{z}(t)$ and wedges $\boldsymbol{\mu}(t)$, given tax–transfer instruments $(\boldsymbol{\tau}, \boldsymbol{T})$. This formulation replaces the inner minimum with a set of inequality constraints and reduces the max–min problem to a linear programming problem that can be solved easily on a computer.

We now describe step-by-step how we calculate aggregate productivity in the specific context of our application.

Step 0: Calibrate the 10-country, 30-industry, 4-factor-per-industry model to the 2008 trade data (assuming trade costs but no tariffs) and set the elasticities at the values described in the text.

Step 1: Feed into the model small increments in U.S. tariffs (and retaliatory tariffs from the rest of the world). Stop raising tariffs when the total quantity of the consumption good in the U.S. is maximized. The resulting allocation is the status quo. The status quo tariffs are 1% with factor mobility and 2.3% without factor mobility.

Step 2: Feed into the model an increase in China's factor endowments so that China accounts for 17.5% of world GDP.

Step 3: Calculate $\Delta \log A$ under different scenarios, with and without factor mobility.

- **Laissez-faire (no tariffs or redistribution policy)** $\Delta \log A^{\text{costly}}$ is the minimum increase in real factor payments across U.S. factors (log change, relative to status quo, in nominal factor payments deflated by the consumption price index in the U.S. across all U.S. factors).
- **Tariffs & non-targeted rebates**
  (a) Given unchanged tariffs, compute $\Delta \log A^{\text{costly}}$ as the minimum increase (relative to status quo) in real factor payments across U.S. factors.
  (b) Feed small increment in U.S. tariff and retaliatory tariffs from the rest of the world.
  (c) Rebate new tariff revenues to factors according to each factor's share in total payments in the status quo.
  (d) Compute minimum increase (relative to status quo) in real factor payments across U.S. factors, and compare with $\Delta \log A^{\text{costly}}$ in (a). If the objective decreased, stop and consider the previous iteration as the final one. Otherwise, return to (b).
- **Tariffs & targeted rebates:**
  - Same procedure as previous case, but in step (c) the new tariff revenue is redistributed to the factor (or factors) with the minimum increase in real factor payments across U.S. factors. Assignment of tariff revenues, conditional on the value of the tariff, is a small linear programming problem.
- **Tariffs & lump-sum transfers**
  (a) Feed small increments in U.S. tariffs and retaliatory tariff from the rest of the world until total quantity of the consumption good in the U.S. is maximized.
  (b) Use lump-sum transfers to equalize change (relative to status quo) in real factor payments across U.S. factors (this is also equal to the change in the total quantify of the consumption good in the U.S.). This change is equal to $\Delta \log A^{\text{lump sum}}$.