NBER WORKING PAPER SERIES

WELFARE AND OUTPUT WITH INCOME EFFECTS AND TASTE SHOCKS

David Baqaee
Ariel Burstein

Working Paper 28754
http://www.nber.org/papers/w28754

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2021, Revised December 2021

We thank Conor Foley and Sihwan Yang for outstanding research assistance. We thank Fernando Alvarez, Andy Atkeson, Natalie Bau, Paco Buera, Javier Cravino, Joe Kaboski, Greg Kaplan, Stephen Redding, Esteban Rossi-Hansberg, Pierre Sarte, DavidWeinstein, and Jon Vogel for helpful comments. We are grateful to Emmanuel Farhi and Seamus Hogan, both of whom passed away tragically before this paper was written, for their insights and earlier conversations on these topics. This paper received support from NSF grant No. 1947611. The conclusions and analysis are our own, calculated in part on data from Nielsen Consumer LLC and provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Welfare and Output with Income Effects and Taste Shocks
David Baqaee and Ariel Burstein
NBER Working Paper No. 28754
May 2021, Revised December 2021
JEL No. E0,E01,E1,E21,E23,E3

ABSTRACT

We characterize how welfare responds to changes in budget sets and technologies when preferences are non-homothetic or subject to shocks, in both partial and general equilibrium. We generalize Hulten's theorem, the basis for constructing aggregate quantity indices, to this context. We show how to calculate the response of welfare to a shock using only knowledge of expenditure shares and elasticities of substitution (and not of income elasticities and taste shocks). We also characterize the gap between welfare and chain-weighted indices. We apply our results to long- and short-run phenomena. In the long-run, we show that if structural transformation is caused by income effects or changes in tastes, rather than substitution effects, then Baumol's cost disease is twice as important for our preferred measure of welfare (equivalent variation at final preferences). In the short-run, we show that standard deflators understate welfarerelevant inflation because product-level demand shocks are positively correlated with price changes. Finally, using the Covid-19 recession we illustrate the differences between partial and general equilibrium notions of welfare, and show that real consumption and real GDP are unreliable metrics for measuring welfare or production.

David Baqaee
Department of Economics
University of California at Los Angeles
Bunche Hall
Los Angeles, CA 90095
and CEPR
and also NBER
baqaee@econ.ucla.edu

Ariel Burstein
Department of Economics
Bunche Hall 8365
Box 951477
UCLA
Los Angeles, CA 90095-1477
and NBER
arielb@econ.ucla.edu

1 Introduction

In this paper, we study how a change in the economic environment affects welfare. For example, how does an individual's welfare change when her budget constraint changes, or how does national welfare change when technologies change? In practice, chain-weighted real consumption, measured by statistical agencies around the world, is often used to answer questions like these.[1] However, the theoretical justification for the use of these measures requires that preferences be homothetic and that there be no taste shocks. These assumptions are highly convenient, but also highly counterfactual. In this paper, we relax both of these assumptions and characterize changes in welfare, changes in chained-weighted consumption, and the gap between the two in terms of measurable sufficient statistics.

To measure welfare, we use a money-metric measure of welfare for final (post-shock) preferences. That is, to compare a choice set at t_0 to one at t_1, we ask: "*under t_1 preferences, how much must consumers' initial endowment change to make them indifferent between their choice sets at t_0 and t_1?*" where t typically refers to time.[2] We focus on this measure, which is an equivalent variation, because it is a money-metric representation of utility (as opposed to compensating variation, which is not a money-metric, see McKenzie and Pearce, 1982). We also focus on final preferences (as opposed to initial preferences) because final preferences are more policy-relevant for temporal comparisons (see Fisher and Shell, 1968). While our focus is on equivalent variation at final preferences, our results can also be used to study other welfare questions (i.e. compensating variation and initial preferences).

We begin by studying this problem in partial equilibrium, where our welfare metric compares and ranks different budget constraints. We then propose a generalization of money-metric measures that ranks production possibility frontiers (PPFs) rather than budget lines. Whereas the partial equilibrium problem asks a microeconomic question — comparing two budget sets for an infinitesimal agent who does not alter market-level prices through her choices — the general equilibrium problem asks a macroeconomic question accounting for the fact that prices are endogenous to choices. Comparing budget constraints (micro welfare) is not equivalent to comparing PPFs (macro welfare) unless preferences are

[1] Chain-weighted indices weigh changes in prices by good-specific weights that are updated every period. As compared to fixed-weight indices, chain-weighted indices account for substitution by consumers. The continuous time analog to a chain-weighted index is called a *Divisia* index. Chained-weighted indices are used to calculate most measures of real economic activity and price deflators, ranging from aggregates like output (real GDP), total factor productivity, private consumption and investment, to less aggregated objects like industry-level measures of production and inflation. The fact that under suitable assumptions these indices approximately measure changes in welfare and production justifies their recommended use in the United Nations' System of National Accounts (see e.g. Chapters 15 and 17 in IMF, 2004).

[2] Although in this paper we treat t as an index of time, in principle t can also index location.

homothetic and stable or the PPFs are linear (in which case relative prices do not depend on demand).

In partial equilibrium, measuring welfare requires integrating Hicksian (or compensated) demand curves with respect to prices. We prove that there exists a general equilibrium counterpart to Hicksian demand, and the integral of this object yields welfare changes due to changes in technology in general equilibrium, in an exact parallel to the partial equilibrium theory. This general equilibrium integral also generalizes Hulten's theorem to measure welfare in environments featuring taste shocks and non-homotheticities. For simplicity, in the body of the paper, we focus on neoclassical economies with homogeneous agents, but this result continues to apply in economies with heterogeneous agents and distortions.

We provide exact and approximate characterizations of the change in micro and macro welfare. In contrast to a chain-weighted consumption index, which weighs changes in prices or technologies using observed expenditure shares, welfare-relevant indices weigh changes in prices or technologies using Hicksian expenditure shares. This implies that, compared to our baseline welfare measure, chain-weighted consumption *undercounts* expenditure-switching due to income effects or taste shocks (but not substitution effects).

To understand why chain-weighted indices undercount expenditure-switching due to income effects and taste shocks, consider the following example. Over the post-war period, spending on healthcare grew relative to manufacturing. Suppose this was caused by consumers getting older and richer, because older and richer consumers spend more on healthcare. In this case, a chain-weighted consumption index does not correctly account for expenditure-switching by consumers. Intuitively, when we compare the past to the present, we must use demand curves that are relevant for the older and richer consumers of today, and not demand curves that were relevant in the past. Whereas a chained deflator weighs changes in prices that happened during the 1950's using demand from the 1950's, a welfare-relevant index uses demand from today to weigh changes in prices throughout the sample. We show that the chain-weighted consumption index is higher than the welfare-relevant index if income- or taste-driven expenditure-switching is correlated with changes in prices.

Our results for welfare and the gap between welfare and real consumption are expressed in terms of measurable sufficient statistics. In both partial and general equilibrium, we show that computing the change in welfare does not require direct knowledge of the taste shocks or income elasticities. Instead, what we must know are expenditure shares and elasticities of substitution at the final allocation. For micro welfare, these are the household's expenditure shares and elasticities of substitution. For macro welfare, these are the input-output table and elasticities of substitution in both production and consumption. Our results can be used both for ex-post accounting and ex-ante counterfactuals.

For very simple economies with one factor, constant returns to scale, and no intermediates, the difference between welfare and chain-weighted consumption is approximately half the covariance of supply and demand shocks. We generalize this formula to more complex economies and show how the details of the production structure, like input-output linkages, complementarities in production, and decreasing returns to scale, interact with non-homotheticities and preference shocks and can magnify the gap between welfare and chain-weighted consumption. The discrepancies between welfare and chain-weighted consumption that we emphasize do not get "aggregated" away. In fact, the more we disaggregate, the more important these discrepancies are likely to become. In this sense, our results are related to the literature studying the macroeconomic implications of production networks and disaggregation (e.g. Gabaix, 2011; Acemoglu et al., 2012; Baqaee and Farhi, 2019b).

We illustrate the relevance of our results for understanding long-run and short-run phenomena with three applications.

i. **Long-run application:** Since Baumol (1967), an enduring stylized fact is that industries with slow productivity growth tend to become larger as a share of the economy over time. This phenomenon, known as Baumol's cost disease, implies that aggregate growth is increasingly determined by productivity growth in slow-growth industries since, over time, the industrial mix of the economy shifts to favor these industries. To be specific, from 1947 to 2014, aggregate TFP in the US grew by 60%. If the US economy had kept its original 1947 industrial structure, then aggregate TFP would have grown by 78% instead. We show that if this transformation is caused solely by income effects and demand instability, then our baseline measure of welfare-relevant TFP grew by only 47% instead of 60%.

ii. **Short-run application:** While our first application focuses on long-run patterns, our second application shows that gaps between real consumption and welfare are likely present at high frequencies too. Whereas industry-level sales shares are relatively stable over short-horizons, firm or product-level sales shares are not. In a product-level specification of our model, we show that when products' demand shocks are correlated with their supply shocks, there is a gap between welfare-relevant and measured changes in industry-level output and prices. These biases do not disappear as we aggregate up to the level of real GDP even if products and industries are infinitesimal. When we work with industry-level (rather than disaggregated firm- or product-level) data, we rule out the existence of these biases by assumption. We quantify these biases at the industry level using product-level non-durable consumer goods data between 2004 and 2019. We find that standard price indices, like the Sato-Vartia index and chained-weighted price index, understate the welfare-relevant inflation rate by around 1 percentage point between 2018 and 2019, and this bias grows to 4.3 percentage points over the whole sample.

iii. **Business-cycle application:** Our final application draws on the Covid-19 recession to illustrate the difference between macroeconomic and microeconomic notions of welfare. During this recession, household expenditures switched to favor certain sectors at the same time that those sectors experienced higher inflation. We show that this implies that microeconomic welfare, taking changes in prices as given, fell by more than macroeconomic welfare, taking into account the fact that changes in prices are themselves caused by demand shocks. Furthermore, real consumption failed to measure either object. This is because in episodes where spending patterns are partly driven by taste shocks, as in the Covid-19 recession, changes in real consumption generically depend on irrelevant details like the order in which supply and demand shocks hit the economy. In these circumstances, the change in real consumption between two time periods is not a function of only prices and quantities in those two periods. Real consumption can be different between the initial and final periods even if initial and final prices and quantities of every good stays the same. The same logic applies to real GDP, which means that real GDP or TFP are unreliable metrics for measuring changes in productive capacity in these circumstances.

In addition to preference stability and homotheticity, a chained index accurately measures welfare only if prices and quantities are measured correctly and continuously. Many of the well-known reasons why chained indices fail to measure welfare are due to violations of these measurement assumptions. For example, it is well-known that real consumption fails to account for product creation and destruction if we do not measure the quantity of goods continuously as their price falls from or goes to their choke price (Hicks, 1940; Feenstra, 1994; Hausman, 1996; Aghion et al., 2019); real consumption does not properly account for changes in the quality of goods (see Syverson, 2017); and, real consumption fails to properly account for changes in non-market components of welfare, like changes in the user-cost of durable consumption or leisure and mortality (see Jones and Klenow, 2016). In all of these cases, the problem is that some of the relevant prices or quantities in the consumption bundle are missing or mismeasured, and correcting the index involves imputing a value for these missing prices or quantities. Non-homotheticities and taste shocks are different from mismeasured prices because they violate the maintained assumptions about preferences, not prices, and correcting the index requires the use of welfare-relevant, rather than observed, expenditure shares. For this reason, we abstract from these mismeasurement issues and assume that prices and quantities have been correctly measured. If prices and quantities are mismeasured or missing, then our results would apply to the quality-adjusted, corrected, version of prices instead of observed prices.[3]

Other related literature. Measuring changes in welfare using money-metrics is standard in microeconomic theory (see, e.g. chapter 7 in Deaton and Muellbauer, 1980). We characterize the gap between this notion of welfare and real consumption with non-homotheticities and taste shocks. Our general equilibrium results relax the standard assumption in growth accounting that there exists a stable and homothetic final good aggregator (extending Domar 1961, and Hulten 1978). This is an issue of central importance in the literature on disaggregated and production network models (see, for example, Carvalho and Tahbaz-Salehi, 2018 and the references therein).

Our approach focusing on the money-metric at final preferences can be contrasted with common practice in the literature on index numbers, which focuses on Konüs price indices for intermediate levels of utility or tastes between t_0 and t_1 (see, for example, Diewert 1976, Caves et al. 1982, and Feenstra and Reinsdorf 2007). These papers show that, under some assumptions (i.e. translog or CES), commonly used indices like Tornqvist and Sato-Vartia do answer an economically meaningful question. The advantage of this approach is that constructing these indices requires far less information; the disadvantage is that, unlike our preferred welfare measure, these indices are not money-metrics that can be used for policy or counterfactual analysis, and they do not provide specific information about the reference indifference curve being used (what budget level, price vector, and preferences it corresponds to). Furthermore, in practice, most index numbers are constructed by chaining, and the aforementioned results do not apply to chained indices. An additional contribution of our paper is to characterize how equivalent variation differs from chained (Divisia) indices under arbitrary price and income paths.[4] Finally, relative to this literature,

[3]Our approach to calculate ex-post welfare changes requires well-measured price changes, as well as budget shares and elasticities of substitution in the final period. For ex-post welfare measurement, when information on prices is missing or mismeasured, if preferences are non-homothetic an alternative approach is to infer changes in welfare by relying on changes in prices, expenditures, price elasticities, and Engel curve slopes for only a subset of goods, given assumptions on separability and stability in preferences (see e.g. Hamilton, 2001 and, more recently, Atkin et al., 2020).

[4]Under the assumption that the path of prices is linear, Feenstra and Reinsdorf (2000) shows the equivalence between Divisia and a Konüs price index for an intermediate utility level under AIDS preferences. In practice, price paths tend to be nonlinear (for evidence using scanner-level data, see Ivancic et al., 2011). Therefore, in contrast to Tornqvist and Sato-Vartia, chained indices cannot generically be interpreted as welfare measures corresponding to any well-defined preferences. This is because, as we discuss in Section 5, Divisia (or chained) indices are path-dependent, so they can violate basic properties like assigning a higher value to a strictly larger choice set. Oulton (2008) discusses how Konüs price indices resolve the

we also provide a unified analysis of non-homotheticity and taste shocks, and we define and characterize a welfare measure for comparing PPFs rather than budget sets (taking into account that prices are endogenous to choices).

A recent and related paper is Redding and Weinstein (2020), who also study welfare changes with both taste shocks and non-homotheticities. Whereas their approach to measuring welfare uses cardinal properties of the utility function, we use a money-metric (or endowment-metric in general equilibrium) that relies only on ordinal properties of preference relation. We compare our approach to that of Redding and Weinstein (2020) in detail in Appendices E.2 and F.[5]

Our paper is also related to the literature on structural transformation and Baumol's cost disease. As explained by Buera and Kaboski (2009) and Herrendorf et al. (2013), this literature advances two microfoundations for structural transformation. The first explanation is all about relative prices differences: if demand curves are not unit-price-elastic, then changes in relative prices change expenditure shares (e.g. Ngai and Pissarides, 2007; Acemoglu and Guerrieri, 2008; Buera et al., 2015). The second explanation emphasizes shifts in demand curves caused by income effects or taste shocks— households spend more of their income on some goods as they become richer (e.g. Kongsamut et al., 2001; Boppart, 2014; Comin et al., 2015; Alder et al., 2019) or older (Cravino et al., 2019). Our results suggest that structural transformation driven by relative price changes has different welfare implications than structural transformation driven by non-homotheticity or taste shocks.

The structure of the paper is as follows. In Section 2, we set up the microeconomic problem and provide exact and approximate characterizations of the difference between welfare and measured real consumption. In Section 3, we set up the macroeconomic general equilibrium model and provide exact and approximate characterizations of the difference between welfare and measured real output changes. Whereas in Section 3 we present our macro results in terms of endogenous sufficient statistics, in Section 4 we solve for these endogenous sufficient statistics in terms of microeconomic primitives and consider some simple but instructive analytical examples. Our applications are in Section 5. We discuss some extensions, including the treatment of new goods, distortions, and household heterogeneity, in Section 6 and conclude in Section 7.

---

path-dependency problem of Divisia indices.

[5]For background on how to account for taste shocks in welfare measures, see Fisher and Shell (1968) and Samuelson and Swamy (1974). For a comparison of measures of welfare that rely on cardinal and ordinal properties of utility in settings with changes in tastes, see Muellbauer (1975), Balk (1989), and Martin (2020).

2 Microeconomic Changes in Welfare and Consumption

In this section, we consider changes in budget constraints in partial equilibrium. We ask how consumers value these changes, and compare this with chain-weighted real consumption. This section builds intuition for Section 3, where we model the equilibrium determination of prices.

2.1 Definition of Welfare and Real Consumption

Consider a set of preference relations, $\{\succeq_x\}$, over bundles of goods $c \in \mathbb{R}^N$, where N is the number of goods. The vector c includes all relevant goods, and if $\succeq_x$ is intertemporal, then c is a path of current and future consumption bundles.[6]

These preferences are indexed by some parameter x that the consumer does not make choices about but that can affect preference rankings over bundles of goods. For example, x could be calendar time, age, exposure to fads, or state of nature. For every x, we represent the preference relation $\succeq_x$ by a utility function $u(c;x)$. Since the consumer makes no choices over x, preferences over x, if they exist, are not revealed by choices. Hence, whereas $u(c;x) > u(c';x)$ if, and only if, $c \succeq_x c'$, a comparison of $u(c;x')$ and $u(c;x)$ is not meaningful and does not encode any information, because it is affected by how $u(\cdot;x)$ and $u(\cdot;x')$ are cardinalized.

There are two properties of preferences that are analytically convenient benchmarks throughout the rest of the analysis.

Definition 1. Preferences are *homothetic* if whenever $c \sim_x c'$ then $\alpha c \sim_x \alpha c'$ for every $\alpha > 0$.

When $\succeq_x$ is homothetic, we can write $u(c;x)$ so that for every $\alpha > 0$, $u(\alpha c;x) = \alpha u(c;x)$.

Definition 2. Preferences are *stable* if $\succeq_x$ is the same as $\succeq_{x'}$ for every x and x'.

If preferences are stable, then the utility function $u(c;x)$ is separable in c and x.

The indirect utility function, for any value of x, is

$$v(p, I; x) = \max_c \{u(c;x) : p \cdot c = I\},$$

where p is a price vector over goods and I is expenditures (which we interchangeably refer to as income).

Consider shifts in the budget set as prices and income change from p_{t_0} and I_{t_0} to p_{t_1} and I_{t_1}. Here, t_0 and t_1 simply index the vector of prices and income being compared.

[6] We explicitly discuss how to apply our results in dynamic economies in Section 4.3.

Motivated by our applications, we refer to this index as time. This change in the budget set is accompanied by changes in preferences from x_{t_0} to x_{t_1}.

Since utility is only defined up to monotone transformations, changes in utility do not have meaningful units. Therefore, to measure how the consumer values different budget sets, we use the money-metric in t_0 prices. Our baseline measure of microeconomic welfare is defined as follows.

Definition 3 (Micro Welfare)**.** The change in welfare measured using the *micro equivalent variation* with *final preferences* is $EV^m(p_{t_0}, I_{t_0}, p_{t_1}, I_{t_1}; x_{t_1}) = \phi$ where ϕ solves

$$v(p_{t_1}, I_{t_1}; x_{t_1}) = v(p_{t_0}, e^{\phi} I_{t_0}; x_{t_1}). \tag{1}$$

In words, EV^m is the change in income (in logs), under initial prices p_{t_0}, that a consumer with preferences $\succeq_{x_{t_1}}$ would need to be indifferent between the budget set defined by initial prices $(p_{t_0}, e^{\phi} I_{t_0})$ and the new budget set defined by new prices and income (p_{t_1}, I_{t_1}). The new budget set is preferred to the initial one, if and only if, EV^m is positive.

Discussion of our welfare criterion. Our welfare measure does not attempt to measure the change in income a consumer with t_1 preferences needs to be as well off as a consumer with t_0 preferences. This is because answering such a question is impossible using choice data. Instead, a well-defined welfare measure must hold preference parameters x constant in the comparison, since the preference relation $\succeq_x$ does not encode information about preferences over x itself. That is, a comparison of $u(\cdot; x)$ and $u(\cdot; x')$ is not meaningful because it would depend on the arbitrary choice of cardinalization embedded in the utility function.

Our baseline welfare measure focuses on *final* rather than *initial preferences*, and *equivalent* rather than *compensating* variation. In principle, one could study initial preferences and compensating variation instead. In general, these alternative measures give different answers unless preferences happen to be both stable and homothetic. We briefly discuss in the body of the paper how our results change for different welfare measures, and provide more details in Appendix C. We choose to focus on final preferences $\succeq_{x_{t_1}}$, as opposed to initial preferences $\succeq_{x_{t_0}}$, because for temporal comparisons, the asymmetry of time makes current preferences more relevant than preferences in the past. As Fisher and Shell (1968, page 5) write, "*...every practical question which one wants the cost of living index to answer is answered with reference to current, not base-year tastes.*"

The other choice we make is to use equivalent variation, rather than compensating variation, as our benchmark. We focus on equivalent variation because, unlike the compensating variation, the equivalent variation is a money-metric (see McKenzie and Pearce, 1982).

Specifically, the equivalent variation is itself an index of utility which transforms the utility value of different outcomes into dollar values under a common price system (prices in t_0). In other words, $(p, I) \succeq_x (p', I')$ if, and only if, $EV^m(p_{t_0}, I_{t_0}, p, I; x) \geq EV^m(p_{t_0}, I_{t_0}, p', I'; x)$. This is not true for compensating variation.

We now define changes in chain-weighted real consumption. This corresponds to what national income accountants and statistical agencies do when given data on the evolution of prices p and consumption bundles c.

Definition 4 (Real consumption)**.** For some path of prices and quantities that unfold as a function of time t, the change in *real consumption* from t_0 to t_1 is defined to be

$$\Delta \log Y = \int_{t_0}^{t_1} \sum_{i \in N} b_{it} \frac{d \log c_{it}}{dt} dt = \int_{t_0}^{t_1} \sum_{i \in N} b_i d \log c_i, \tag{2}$$

where $b_{it} \equiv p_{it} c_{it} / I_t$ is the budget share of good i given prices, income, and preferences at time t.[7]

The last equation on the right-hand side suppresses dependence on t in the integral. We sometimes use this convention to simplify notation. Equation (2) is called a *Divisia* quantity index. In practice, since perfect data is not available in continuous time, statistical agencies approximate this integral via a (Riemann) sum using chained indices. We abstract from the imperfections of these discrete time approximations in this paper.[8] Moreover, we assume that the data on prices and quantities is *perfect* — completely accurate, comprehensive, and adjusted for any necessary quality changes. This is because the important and well-studied biases associated with imperfections in the data, like the lack of quality adjustment, missing prices, or infrequent measurement, are different to the biases we study. Appendix F shows that the welfare implications of taste shocks and quality changes are very different.[9]

[7]For any variable z, we denote by dz its change over infinitesimal time intervals, so that $\Delta z = \int_{t_0}^{t_1} dz$.

[8]In discrete time, one can approximate this Riemann integral in different ways. For example, we can use left-Rieman sums (Chained Laspeyres), right-Riemann sums (Chained Paasche), or mid-point Riemann sums (Chained Tornqvist or Fisher). In continuous time, all of these procedures are equivalent and yield the same answer.

[9]The standard approach to modeling quality is hedonics, where goods are a bundle of characteristics and consumers have preferences over characteristics. The BLS uses hedonic adjustment of prices for certain goods in the CPI to account for changing product quality. For example, for computers, CPU speed is a characteristic that consumers value. If a computer increases its CPU speed, the consumer can consume more of this characteristic. Choices made by consumers over computers with different CPU speeds reveal how consumers value this characteristic. Note that there is no reason to normalize the level of quality across goods because the units of characteristics are pinned down (e.g. GHz). However, even after all the quality-adjustments have been done, demand curves can still shift. We model such residual shifts in demand curves as taste shocks and, by definition, they do not involve choices from the consumer's perspective.

Define the expenditure function for any value of x by

$$e(p,u;x) = \min_c \{ \sum_{i \in N} p_i c_i : u(c;x) = u \}.$$

The compensated or Hicksian budget share of good i (given prices p, preferences x, and a level of utility u) is

$$b_i(p,u;x) \equiv \frac{p_i c_i(p,u;x)}{e\,(p,u;x)} = \frac{\partial \log e(p,u;x)}{\partial \log p_i}, \tag{3}$$

where the second equality is Shephard's lemma. Using the budget constraint, real consumption in (2) can be expressed in terms of changes in nominal income deflated by price changes:

$$\Delta \log Y = \Delta \log I - \int_{t_0}^{t_1} \sum_{i \in N} b_i(p_t, u_t, x_t) \frac{d \log p_{it}}{dt} dt, \tag{4}$$

where $b(p_t, u_t, x_t)$ are observed budget shares at t. In words, changes in real consumption are equal to changes in income minus changes in the consumption price deflator. Changes in real consumption (and the consumption price deflator) potentially depend on the entire path of prices and quantities between t_0 and t_1 and not just the initial and final values. This is unlike welfare changes, EV^m, which depend only on initial and final prices and incomes and not on their entire path.

2.2 Relating Welfare and Consumption

We consider how real consumption and welfare change in response to changes in the budget set and the preferences of the consumer. We first consider globally exact results and then local approximations. The results are stated in terms of changes in prices and income, which we endogenize in Sections 3 and 4.

Global results. We start by expressing changes in welfare in terms of changes in prices and expenditure shares.

Lemma 1 (Micro Welfare)**.** *For any smooth path of prices, income, and tastes that unfold as a function of time t, micro welfare changes are given by*

$$EV^m = \Delta \log I - \int_{t_0}^{t_1} \sum_{i \in N} b_i(p_t, u_{t_1}, x_{t_1}) \frac{d \log p_{it}}{dt} dt, \tag{5}$$

where $b_i(p_t, u_{t_1}, x_{t_1}) = b_i(p_t, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})$ denotes Hicksian budget shares at prices p_t fixing final preferences x_{t_1} and final utility $u_{t_1} = v(p_{t_1}, I_{t_1}; x_{t_1})$.

Comparing (5) and (4) clarifies the differences between welfare and real consumption. Real consumption weighs changes in prices at time t by observed budget shares at time t, taking into account expenditure-switching as it happens. In contrast, welfare takes into account expenditure-switching due to income effects and taste shocks from the beginning, weighing changes in prices at time t by $b_i(p_t, u_{t_1}, x_{t_1})$. Intuitively, EV^m depends on budget shares evaluated at final utility (u_{t_1}) and tastes (x_{t_1}), since EV^m adjusts the level of income in t_0 to make consumers with t_1 preferences as well off as they are in t_1. For instance, if welfare increases from t_0 to t_1, consumers must be given more income in t_0 to make them indifferent between t_0 and t_1. As we give consumers more income in t_0, the shape of their indifference curve changes until it mirrors the one in t_1. This means that the shape of the indifference curve relevant for the comparison is the one at t_1.[10,11]

Lemma 1 follows from the observation that EV^m can be re-expressed, using the expenditure function, as

$$EV^m = \log \frac{e\left(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1}\right)}{e\left(p_{t_0}, v(p_{t_0}, I_{t_0}; x_{t_1}); x_{t_1}\right)} = \Delta \log I - \log \frac{e\left(p_{t_1}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1}\right)}{e\left(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1}\right)}, \tag{6}$$

and recognizing that the second term can be written as the integral in (5). The ratio of expenditure functions, holding fixed utility, is called a Konüs (1939) price index. Equation (6) shows that EV^m requires deflating nominal income by a particular choice of the Konüs price index.[12]

To make the notation more compact, denote the *welfare-relevant* Hicksian budget shares in Lemma 1 by

$$b_i^{ev}\left(p_t\right) \equiv b_i(p_t, u_{t_1}, x_{t_1}).$$

We can reinterpret these hypothetical budget shares $b^{ev}(p)$ as corresponding to those of a fictional consumer with homothetic and stable preferences with expenditure function

[10] When there are no taste shocks, real consumption, defined by (2), is a multi-good version of the change in consumer surplus, which is the area under the Marshallian demand curve. Similarly, by equation (5), welfare is the area under a Hicksian demand curve. Hence, in a partial equilibrium context with stable preferences, the gap between real consumption and welfare is also the gap between consumer surplus and welfare, studied by Hausman (1981) and McKenzie and Pearce (1982) amongst others. This equivalence does not hold when preferences are unstable since Marshallian consumer surplus is not the same as chained real consumption.

[11] By definition, EV^m only depends on initial and final prices and income, given t_1 preferences. By the gradient theorem for line integrals, the integral in (5) is path-independent and can be computed under any continuously differentiable path of prices that go from p_{t_0} to p_{t_1}. When comparing EV^m and real consumption, we consider the integral under the realized path of prices over time, which as described in the text is assumed to be available in continuous time.

[12] Equation (5) shows that EV^m requires the use of a Konüs price index using t_1 utility and preferences — this is in contrast to common practice in the index number theory literature, say Diewert (1976), that uses an intermediate level of utility or preferences between u_{t_0} and u_{t_1}. We discuss how our results relate to this alternative approach in Appendix D.

$e^{ev}(p,u) = e(p, u_{t_1}; x_{t_1})\, u$, where $u_{t_1} = v(p_{t_1}, I_{t_1}; x_{t_1})$. This implies that we can calculate changes in welfare given changes in prices based on budget shares $b^{ev}(p)$, *without* needing to know income elasticities or the nature of demand shocks. This is because this fictional consumer has homothetic and stable preferences, so all income elasticities are equal to one and there are no demand shocks. To compute $b^{ev}(p)$, we need to know the terminal budget shares and the terminal elasticities of substitution, as discussed in the following remark.

Remark 1 (Non-homothetic CES preferences)**.** To illustrate how Lemma 1 can be used, consider a non-homothetic CES example as in Hanoch (1975), Comin et al. (2015), Matsuyama (2019), and Fally (2020).[13] For this demand system, the following equation pins down changes in budget shares at time t:[14]

$$d\log b_{it} = \underbrace{[1-\theta_0]\,[d\log p_{it} - \mathbb{E}_{b_t}(d\log p_t)]}_{\text{substitution effects}} + \underbrace{[\varepsilon_{it}-1]\,[d\log I_t - \mathbb{E}_{b_t}(d\log p_t)]}_{\text{income effects}} + \underbrace{d\log x_{it}}_{\text{taste shock}}, \quad (7)$$

where $\mathbb{E}_b(\cdot)$ is a budget-share weighted average. The elasticity θ_0 is the (constant utility) elasticity of substitution across goods and ε_{it} is the income elasticity of good i. The term $d\log x_{it}$ is a demand shifter (i.e. a taste shock), a residual that captures changes in expenditure shares not attributable to changes in income or prices. Note that when ε_{it} is equal to 1 for every i and t, final demand is homothetic, and when x_{it} is constant for all i and t, final demand is stable.

If we know b_{t_1}, we can construct the welfare-relevant budget shares $b^{ev}(p)$ between t_0 and t_1 by iterating on the differential equation

$$d\log b^{ev}_{it} = [1-\theta_0]\,[d\log p_{it} - \mathbb{E}_{b^{ev}_t}(d\log p_t)], \quad (8)$$

starting at t_1 with initial value $b^{ev}_{t_1} = b_{t_1}$ and going back to t_0. These are changes in budget shares which are only due to substitution effects, and hence omit the last two terms in equation (7). Given the path of b^{ev}_t, we can then apply Lemma 1. For non-homothetic CES,

[13]The result that only terminal budget shares and elasticities of substitution are necessary to calculate EV^m is true for arbitrary non-CES functional forms, but since the intuition for the more general case is similar to the CES case, we leave the more general non-parametric results in Appendix K. We use non-homothetic CES in our worked-out examples since it provides a clean separation between substitution elasticities (necessary for computing welfare) and other parameters of the utility function, in contrast to other commonly used non-homothetic demand systems such as PIGL and AIDS. Furthermore, substitution elasticities are symmetric and constant for CES, which also helps keep the examples intuitive.

[14]We assume that all goods have positive expenditure shares. In Section 6 we show how to calculate changes in welfare when preferences are unstable and there are changes in the set of goods purchased by consumers. See Appendix E.1 for a derivation and mapping between ε and $d\log x$ and primitive preference parameters.

the integral in Lemma 1 has a closed form solution

$$EV^m = \Delta \log I - \int_{t_0}^{t_1} \sum_{i \in N} b_i^{ev}(p_t) \frac{d \log p_{it}}{dt} dt = \Delta \log I + \log \left(\sum_i b_{it_1} \left(\frac{p_{it_0}}{p_{it_1}} \right)^{1-\theta_0} \right)^{\frac{1}{1-\theta_0}}. \quad (9)$$

This shows that the income elasticities and taste shocks are not directly required.[15] Equation (9) shows when CES preferences are non-homothetic or unstable, it is the exact hat algebra CES deflator (often referred to as the Lloyd-Moulton index) from t_1 to t_0 rather than the Sato-Vartia index that must be used.[16]

If b_{t_1} is not known, we first have to predict b_{t_1} by iterating on equation (7) from t_0 to t_1 to obtain b_{t_1}. This first step requires full knowledge of demand shocks and income elasticities over time (see Appendix E for more details). Once in possession of b_{t_1}, apply (9) to get the change in welfare.

Remark 2 (Compensating Variation under Initial Preferences)**.** Our baseline measure of welfare changes is equivalent variation under final preferences. An alternative would be to use compensating variation under initial preferences. Every result in the paper can be translated into compensating variation under initial preferences simply by reversing the flow of time. In particular, whereas Lemma 1 preserves the shape of the indifference curve at the final allocation, the compensating variation counterpart to Lemma 1 preserves the shape of the indifference curve at the initial allocation. Hence, calculating compensating variation requires knowledge of initial budget shares and elasticities of substitution, whereas equivalent variation requires knowledge of final budget shares and elasticities of substitution.[17]

Compensating variation at initial preferences differs from equivalent variation at final preferences (unless preferences are stable and homothetic) because they answer different questions. Suppose that the household ages and becomes richer from t_0 to t_1. Intuitively, the way a rich, old consumer's welfare is affected by price changes is different to the way a young, poor consumer's welfare is affected by price changes. The richer or older household

[15] In practice, estimating the elasticity of substitution θ_0 must take into account the possibility of demand shocks and income effects (via Slutsky's equation). For example, Auer et al. (2021) estimate compensated price elasticities and apply Lemma 1 to measure the heterogeneous welfare effects of changes in foreign prices in the presence of demand non-homotheticities.

[16] A common approach in the literature is to use a price index defined by $P_t \equiv e(p_t, u_t; x_t)/u_t$, where u_t is the utility index under non-homothetic CES preferences. In Appendix E.2, we show that changes in P_t depend on cardinal properties of the utility function that are not pinned down by the ordinal preference relation and hence cannot be disciplined by observable choice data, even when preferences are non-homothetic but stable.

[17] This means that calculating equivalent variation at final preferences is more convenient for ex-post comparisons and compensating variation at initial preferences is more convenient for ex-ante comparisons or counterfactuals.

will place more value on the price of luxuries (non-homotheticity) or healthcare (instability of preferences) than the younger or poorer household. Hence, changes in the relative price of luxuries or healthcare will have different effects depending on whether one uses equivalent or compensating variation, and initial or final preferences.[18]

We now contrast changes in real consumption and welfare.

Proposition 1 (Consumption vs. Welfare). *Given a smooth path of prices, income, and tastes that unfold as a function of time t, the difference between welfare changes and real consumption is*

$$EV^m - \Delta \log Y = \int_{t_0}^{t_1} \sum_{i \in N} \left(b_{it} - b_{it}^{ev} \right) \frac{d \log p_{it}}{dt} dt = (t_1 - t_0) \, \mathbb{E}_t Cov \left(b_t - b_t^{ev}, d \log p_t \right),$$

where the covariance is calculated across goods at a point in time, and the average is calculated across time between t_0 and t_1.

An immediate consequence of Proposition 1 is the well-known result that real consumption is equal to changes in equivalent variation if, and only if, preferences are homothetic and stable. This is because when preferences are stable and homothetic, budget shares do not depend on x or on utility u over time. Hence, whenever preferences are homothetic and stable, $b_{it}^{ev} = b_{it}$ for every path of shocks and every t.

To gain more intuition for the gap between welfare and real consumption, we use a second-order approximation.

Local results. Consider local approximations of the objects of interest as the time period goes to zero, $t_1 - t_0 = \Delta t \to 0$.[19] Around t_0, the change in real consumption is approximately

$$\Delta \log Y \approx \underbrace{\Delta \log I - \mathbb{E}_b(\Delta \log p)}_{\text{first-order}} - \underbrace{\frac{1}{2} Cov_b \left(\Delta \log b, \Delta \log p \right)}_{\text{second-order}}, \tag{10}$$

where $\mathbb{E}_b(\cdot)$ and $Cov_b(\cdot)$ are evaluated using budget shares at t_0 as probability weights. The first-order term is just the change in nominal income deflated by average price changes. The second-order terms depend on how expenditures change in response to the shock, and

[18] In Appendix C we show that, up to a second-order approximation (but not globally), changes in real consumption equal a simple average of equivalent variation under final preferences and compensating variation under initial preferences. This means that, up to a second-order approximation, the gap between real consumption and equivalent variation at final preferences is exactly the negative of the gap between real consumption and compensating variation at initial preferences.

[19] For our local approximations, we assume that the exogenous parameters (prices, income, and taste shifters) are smooth functions of t and that the expenditure function is a smooth function of x.

these changes in expenditures can be driven by either substitution effects, income effects, or taste shocks.

To make the relationship between real consumption and welfare more concrete, we use the non-homothetic CES aggregator introduced in Remark 1.[20]

Proposition 2 (Approximate Real Consumption and Micro Welfare)**.** *Consider some perturbation in tastes $\Delta \log x$, prices $\Delta \log p$, and income $\Delta \log I$. Then, to a second-order approximation, the change in real consumption is*

$$\Delta \log Y \approx \Delta \log I - \mathbb{E}_b\left(\Delta \log p\right) - \underbrace{\frac{1}{2}(1-\theta_0) Var_b\left(\Delta \log p\right)}_{\substack{\text{expenditure-switching} \\ \text{due to substitution effect}}} \tag{11}$$
$$- \frac{1}{2} \underbrace{Cov_b\left(\Delta \log x, \Delta \log p\right)}_{\substack{\text{expenditure-switching} \\ \text{due to taste shock}}} - \frac{1}{2} \underbrace{\left[\Delta \log I - \mathbb{E}_b\left(\Delta \log p\right)\right] Cov_b\left(\varepsilon, \Delta \log p\right)}_{\substack{\text{expenditure-switching} \\ \text{due to income effect}}},$$

and the change in welfare is

$$\begin{aligned} EV^m \approx\ & \Delta \log I - \mathbb{E}_b\left(\Delta \log p\right) - \frac{1}{2}(1-\theta_0) Var_b\left(\Delta \log p\right) \\ & - Cov_b\left(\Delta \log x, \Delta \log p\right) - \left[\Delta \log I - \mathbb{E}_b\left(\Delta \log p\right)\right] Cov_b\left(\varepsilon, \Delta \log p\right), \end{aligned} \tag{12}$$

where $\mathbb{E}_b(\cdot)$, $Var_b(\cdot)$, and $Cov_b(\cdot)$ are evaluated using budget shares at t_0 as probability weights.

Consider the change in real consumption in (11), which rewrites the nonlinear terms in (10) in terms of primitives. Since these are second-order, they are multiplied by 1/2. We discuss these terms one-by-one. If goods are substitutes, $\theta_0 > 1$, then welfare is convex in prices and variance in price changes boosts welfare by raising the expenditure share of goods that become relatively cheap. The second line of (11) captures the effect of taste shocks and income effects. If the composition of demand shifts in favor of goods that become relatively cheap, either due to taste shocks $Cov_b(\Delta \log x, \Delta \log p) < 0$ or income effects $Cov_b\left(\varepsilon, \Delta \log p\right)\left(\Delta \log I - \mathbb{E}_b\left(\Delta \log p\right)\right) < 0$, then real consumption increases.

Now consider changes in welfare in (12). The first-order terms are identical to real consumption, but discrepancies are present at the second order. In particular, welfare places a larger weight on changes in expenditure shares that occurred due to income effects and taste shocks. Whereas $\Delta \log Y$ only takes into account expenditure-switching as it occurs over time, EV^m accounts expenditure-switching due to income effects and taste shocks

[20] For a more elaborate discussion of Proposition 2 without imposing non-homothetic CES, see Proposition 11 in Appendix A. The intuition remains similar.

from the start. Therefore, expenditure-switching due to income and tastes are multiplied by 1 for EV^m and $1/2$ for $\Delta \log Y$. This implies that, for example, the increase in welfare from a price reduction in a good i with increasing demand (due to an increase in x_i or a relatively high ε_i) is not fully reflected in real consumption, implying $EV^m > \Delta \log Y$. If preferences are stable and homothetic, then welfare changes coincide with changes in real consumption. Furthermore, even if preferences are unstable or non-homothetic, real consumption strays from welfare only when price changes covary with non-price changes in demand.[21,22]

In Appendix F we extend Proposition 2 to incorporate unobserved changes in quality. We show that the biases causes by taste shocks are very different to the ones caused by quality changes.

3 Macroeconomic Changes in Welfare and Consumption

In the previous section we showed how to value changes in budget sets, given household preferences. For these problems, the frontier of the consumer's choice set is linear, since prices are assumed to be exogenous. At the level of a whole society however, choice sets need not be linear. The production possibility set associated with an economy may have a nonlinear frontier. In this case, relative prices respond endogenously to choices made by consumers. In this section, we extend our analysis to show how to assign value to different production possibility frontiers (PPFs).

We first generalize our definitions of welfare, now at the macroeconomic level, and we introduce some basic structure and notation. We then present expressions for real GDP and welfare at the macroeconomic level, first globally and then locally in terms of endogenous sufficient statistics. In the next section, Section 4, we solve for these endogenous objects in terms of observable primitives.

[21] In Remark 1 we pointed out that, starting at b_{t_1}, computing welfare does not require knowledge of income elasticities or taste shocks if we know the elasticities of substitution. However, the approximation in (12) depends on income elasticities and taste shocks. The reason is because this approximation is around initial budget shares b_{t_0}. If we start with budget shares at t_1, we get

$$EV^m \approx \Delta \log I - \mathbb{E}_b\left(\Delta \log p\right) + \frac{1}{2}(1-\theta_0)Var_b\left(\Delta \log p\right),$$

where $\mathbb{E}_b(\cdot)$ and $Var_b(\cdot)$ are evaluated using budget shares at t_1. Hence, starting at the terminal budget shares, EV^m only depends on substitution effects as in Remark 1. Both expressions are valid second-order approximations and in either case, real consumption undercounts expenditure-switching caused by income effects or taste shocks.

[22] Whereas EV^m puts more weight than real consumption on expenditure switching due to income effects and taste shocks, compensating variation at initial preferences does the opposite. In particular, to a second order starting at t_0, $CV^m \approx \Delta \log I - \mathbb{E}_b\left(\Delta \log p\right) - \frac{1}{2}(1-\theta_0)Var_b\left(\Delta \log p\right)$.

3.1 Environment and Definitions

Consider a perfectly competitive neoclassical closed economy with a representative agent.[23] Each good $i \in N$ has a production function

$$y_i = A_i G_i \left(\{m_{ij}\}_{j \in N}, \{l_{if}\}_{f \in F} \right),$$

where m_{ij} are intermediate inputs used by i and produced by j, and l_{if} denotes primary factor inputs used by i for each factor $f \in F$. The exogenous scalar A_i is a Hicks-neutral productivity shifter. Without loss of generality, we assume that G_i has constant returns to scale since decreasing returns to scale can be captured by adding producer-specific factors. Furthermore A_i is Hicks-neutral without loss of generality. This is because we can capture non-neutral (biased) productivity shocks to input j for producer i by introducing a fictitious producer that buys from j and sells to i with a linear technology. A Hicks-neutral shock to this fictitious producer is equivalent to a non-neutral technology shock to i.

Let A be the $N \times 1$ vector of technology shifters and L be the $F \times 1$ vector of primary (exogenously given) factor endowments.[24] The production possibility set (and its associated frontier) is the set of feasible consumption bundles that can be attained given A and L. Given our assumption that production functions have constant returns to scale, the PPF is linear if there is only one factor of production.

For each A, L, and x, we denote equilibrium prices and aggregate income by $p(A, L, x)$ and $I(A, L, x)$. These equilibrium prices and incomes are unique up to the choice of a numeraire.

Define the *macro indirect utility* function as the solution to the following planning problem:

$$V(A, L; x) = \max_c \{u(c; x) : c \text{ is feasible}\}.$$

This is the maximum amount of utility the economy can deliver given technologies (A, L) and preferences $\succeq_x$. Whereas the micro indirect utility takes prices as given and lets consumers pick any point in their budget set (even if such a point is not feasible at the economy-wide level), the macro indirect utility function takes the PPF as the primitive and lets society only pick feasible points in the production possibility set. The first welfare theorem implies that the competitive equilibrium decentralizes the planning problem above with prices determined in equilibrium.

[23]For expositional simplicity, we use the representative (or homogeneous) agent assumption and abstract from inefficiencies. See Section 6 for details on how our results can be generalized to economies with heterogeneous agents and distortions.

[24]Allowing for endogenous labor-leisure choice requires including the time endowment in L and leisure in the consumption bundle c.

Consider shifts in the PPF as technologies and factor endowments change from (A_{t_0}, L_{t_0}) to (A_{t_1}, L_{t_1}), along with changes in preferences from x_{t_0} to x_{t_1}. We generalize our microeconomic measure of welfare in the following way.

Definition 5 (Macro Welfare)**.** The change in welfare measured using the *macro equivalent variation* with *final preferences* is $EV^M(A_{t_0}, L_{t_0}, A_{t_1}, L_{t_1}; x_{t_1}) = \phi$ where ϕ solves

$$V(A_{t_1}, L_{t_1}; x_{t_1}) = V(A_{t_0}, e^{\phi} L_{t_0}; x_{t_1}).$$

The superscript M in EV^M represents the fact that this is the *macro* equivalent variation, in contrast to EV^m for the *micro* welfare measure. In words, EV^M is the proportional change in initial factor endowments required to make the consumer with preferences $\succeq_{x_{t_1}}$ indifferent between the PPF defined by $(A_{t_0}, e^{\phi} L_{t_0})$ and the new PPF, defined by (A_{t_1}, L_{t_1}).

Intuitively, EV^M expresses utility changes in terms of factor endowments. That is, one PPF is preferred to another if, and only if, EV^M for the first PPF is higher than the other. Furthermore, EV^M is itself an index of utility in the sense that $(A, L) \succeq_x (A', L')$, if and only if, $EV^M(A_{t_0}, L_{t_0}, A, L; x) \geq EV^M(A_{t_0}, L_{t_0}, A', L', x)$. The macro equivalent variation, EV^M, is a useful metric because it ranks PPFs without reference to endogenous prices. In this sense, EV^M is similar to consumption-equivalents commonly used to measure welfare in macroeconomics.[25] For a graphical representation of EV^m and EV^M, see Figure 1. As we discuss below, EV^M and EV^m always coincide if preferences are homothetic and stable or if the PPF is linear. Moreover, as we show below, measuring welfare changes in terms of factor endowments results in a Hulten-type expression for EV^M.

Macro EV^M and micro EV^m welfare changes are different because they answer different questions. For example, consider a situation where households age between t_0 and t_1 but technologies and factor endowments stay the same. Since the PPF is unchanged, the change in macro welfare is zero by construction. However, if the PPF is nonlinear, the relative price of goods changes between t_0 and t_1: prices rise for those goods that become more desirable. In this case, EV^m necessarily falls even though the PPF is unchanged. Hence, EV^m does not rank PPFs for a society (for more details, see Example 3 below).

The issue is that using the initial budget set to represent the initial PPF is deceptive, since the initial budget set reflects both the technologies and demand in t_0. Our macroeconomic notion of welfare accounts for the endogenous changes in prices by comparing the initial and final PPFs rather than the initial and final budget sets. To compare initial and final PPFs, we scale factor endowments instead of the nominal income endowment,

[25] When preferences are stable and homothetic, EV^M is the same as consumption equivalents, but we do not define welfare changes in terms of consumption equivalents because when preferences are non-homothetic or unstable, households' desired consumption bundle is not stable.

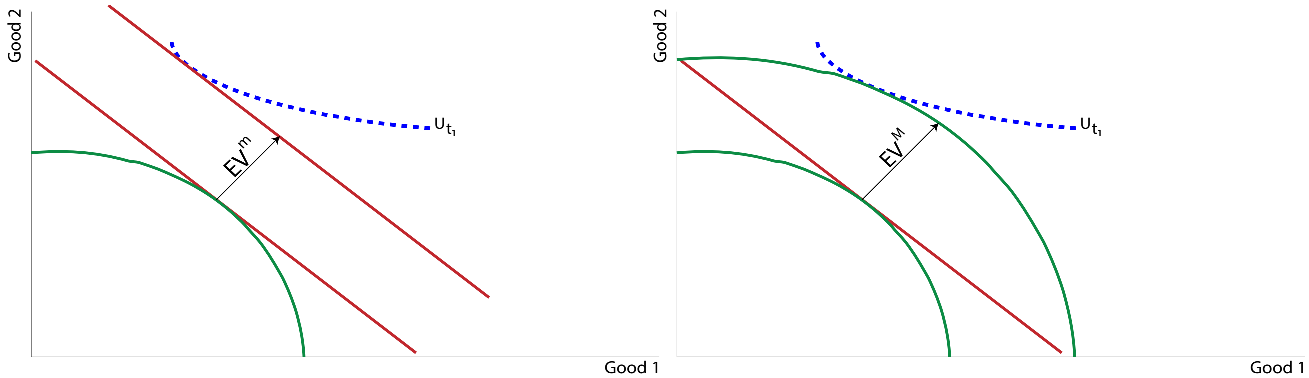


Figure 1: The left panel shows the change in t_0 income that makes the household with final preferences indifferent between the budget constraint in t_0 and t_1. The right panel shows the change in t_0 endowments that makes the household with final preferences indifferent between the PPF in t_0 and t_1. EV^m and EV^M are equal if the PPF is linear, or if preferences are homothetic and stable.

since a proportional shift in factor quantities results in a linear expansion of the PPF and is interpretable without reference to base prices.

When relative prices do not respond to consumers' choices (i.e. the PPF is linear), then for a given primitive shock, macro and micro welfare are always the same. Alternatively, if preferences are homothetic and stable, then macro and micro welfare are the same (regardless of the shape of the PPF). The following proposition formalizes this:

Proposition 3 (Macro vs. Micro Welfare)**.** *Consider changes in technologies A, factor quantities L, and tastes x. Macro and micro welfare changes are equal ($EV^m = EV^M$) if preferences are stable and homothetic, or if factor income shares are constant (i.e. the PPF is linear).*

For a quantitative illustration of the difference between micro and macro welfare see the Covid-19 case study in Section 5.

3.2 Relating Welfare and Real GDP

We now characterize changes in real GDP and welfare, first globally and then locally. The results in this subsection are the general equilibrium counterparts to those in Section 2. They are "reduced-form" in the sense that they are not expressed in terms of primitives. In Section 4, we explicitly solve for these sufficient statistics in terms of observable primitives.

As in Section 2, to study this problem we index the path of technologies, factor endowments, and preferences by time t. The definition of $\Delta \log Y$ is the same as before: $\Delta \log Y = \int_{t_0}^{t_1} \sum_{i \in N} b_{it} d \log c_{it}$. In the general equilibrium model and (its applications), we refer to $\Delta \log Y$ as real GDP.

Denote the sales shares relative to GDP of each good or factor i by

$$\lambda_i = \frac{p_i y_i}{I}\mathbf{1}(i \in N) + \frac{w_i L_i}{I}\mathbf{1}(i \in F),$$

where $\mathbf{1}$ is an indicator function, and w_i and L_i are the price and quantity of factor i. The sales share λ_i is often referred to as a *Domar* weight. Note that referring to λ_i as a "share" is an abuse of language since $\sum_{i\in N} \lambda_i > 1$ whenever there are intermediate inputs.

In equilibrium, the sales shares λ are a function of the state of the economy (A, L, x). To characterize welfare and compare it to real GDP, we introduce *compensated* or *Hicksian* sales shares, which are functions of the PPF, indexed by (A, L), and a specific indifference curve, indexed by u and x.

Definition 6 (Hicksian Sales Shares)**.** The Hicksian (or compensated) sales shares $\lambda(A, L, u, x)$ are sales shares in a fictional economy with the PPF (A, L) but where consumers have stable homothetic preferences represented by the expenditure function $\tilde{e}(p, \tilde{u}) = e(p, u; x)\tilde{u}$.

Global Results. The following results show that changes in real GDP and welfare can be recovered by integrating observed and compensated sales shares with respect to technology changes.

Proposition 4 (Real GDP)**.** *Given a path of technologies, factor quantities, and tastes that unfold as a function of time t, the change in real GDP is*

$$\Delta \log Y = \int_{t_0}^{t_1} \sum_{i\in N} \lambda_i(A_t, L_t, u_t, x_t) \frac{d \log A_{it}}{dt} dt + \int_{t_0}^{t_1} \sum_{i\in F} \lambda_i(A_t, L_t, u_t, x_t) \frac{d \log L_{it}}{dt} dt, \tag{13}$$

where $\lambda(A_t, L_t, u_t, x_t)$ are observed sales shares at t.

In (13), the first N summands are equal to measured TFP, and the last F summands are the growth in real GDP caused by changes in factor endowments. Proposition 4 is a slight generalization of Hulten (1978) to environments with unstable and non-homothetic final demand. In general equilibrium, the sales shares play the role that budget shares played in partial equilibrium. Whereas in partial equilibrium, integrating budget shares with respect to prices yielded real consumption, in general equilibrium integrating sales shares with respect to technologies and factors yields real GDP.

The distinction between macro welfare and real GDP can be understood in terms of differences between observed and Hicksian sales shares.

Proposition 5 (Macro Welfare)**.** *For any smooth path of technologies, factor quantities, and tastes that unfold as a function of time t, changes in macro welfare are*

$$EV^M = \int_{t_0}^{t_1} \sum_{i \in N} \lambda_i(A_t, L_t, u_{t_1}, x_{t_1}) \frac{d \log A_{it}}{dt} dt + \int_{t_0}^{t_1} \sum_{i \in F} \lambda_i(A_t, L_t, u_{t_1}, x_{t_1}) \frac{d \log L_{it}}{dt} dt. \quad (14)$$

In words, growth accounting for welfare should be based on compensated or Hicksian sales shares evaluated at current technology but for final preferences and utility. Analogously to real GDP, we define the first N summands of (14) to be changes in *welfare-relevant TFP* and the last F summands are changes in welfare due to changes in factor inputs. We discuss two salient implications of this proposition.

The first implication is that for welfare questions, the only information we need about preferences are expenditure shares and elasticities of substitution at the final allocation, since the fictional consumer in Proposition 5 has stable preferences with income elasticities all equal to one.[26]

Second, if the path of technologies and factor quantities is continuously differentiable, then real GDP is equal to the change in welfare if, and only if, preferences are homothetic and stable, in which case $\lambda(A, L, u_{t_1}, x_{t_1}) = \lambda(A, L, u_t, x_t)$.

To get more intuition for Propositions 4 and 5, in the following section we use a second-order approximation to characterize changes in real GDP and welfare.

Local Results. We characterize, up to a second order approximation (as $t_1 - t_0 = \Delta t \to 0$), the response of real GDP and welfare to technology and preference shocks, now taking into account the endogenous evolution of sales shares. To make the formulas more compact and without loss of generality, when we write local approximations we abstract from shocks to factor endowments ($\Delta \log L = 0$).[27]

Proposition 6 (Approximate Real GDP and Macro Welfare)**.** *Up to a second order approximation, the change in real GDP is*

$$\Delta \log Y \approx \sum_{i \in N} \lambda_i \Delta \log A_i + \frac{1}{2} \sum_{i \in N} \Delta \lambda_i \Delta \log A_i, \quad (15)$$

[26]Following the observation made in Remark 2, for compensating variation at initial preferences, we need to know elasticities of substitution at the initial allocation instead of the final one.

[27]Shocks to factor endowments are a special case of TFP shocks. To represent a factor endowment shock as a TFP shock, we add fictitious producers that buy the factor endowments on behalf of the other producers and shock their productivity.

and the change in welfare is

$$EV^M \approx \Delta \log Y + \frac{1}{2} \underbrace{\sum_{i \in N} \left[\Delta \log x' \frac{\partial \lambda_i}{\partial \log x} + \Delta \log V \frac{\partial \lambda_i}{\partial \log u} \right] \Delta \log A_i}_{\text{expenditure-switching due to taste shocks and income effects}} . \quad (16)$$

We discuss (15) and (16) in turn. The first term in (15) is the Hulten-Domar formula. The second term captures nonlinearities due to changes in sales shares (since these are second-order, they are multiplied by 1/2). Intuitively, if sales shares decrease for those goods with higher productivity growth, then real GDP growth slows down. This type of effect, known as Baumol's cost disease, is an important driver of the slow-down in aggregate productivity growth.

Equation (16) shows that the gap between macro welfare and real GDP is similar to that for our micro results (the signs are flipped because a positive productivity shock reduces prices). Specifically, real GDP takes into consideration changes in sales shares *along* the equilibrium path. These changes in sales shares could be induced by technology shocks but they could also be due to changes in preferences and non-homotheticities. However, welfare treats changes in shares due to technology shocks differently than changes in shares due to demand shocks or non-homotheticities. That is, real GDP "undercorrects" for changes in shares caused by non-homotheticities or changes in preferences. These terms are multiplied by 1/2 in real GDP, but they are multiplied by 1 in welfare, similar to the partial equilibrium counterpart in Proposition 2.

4 Structural Macro Results and Analytic Examples

The results in Section 3 are reduced-form in the sense that they take changes in observed and compensated sales shares as given. In this section, we solve for changes in these endogenous objects in terms of observable sufficient statistics. For clarity, we restrict attention to nested-CES economies. The general case is in Appendix K, and the intuition is very similar. After providing a characterization to solve for changes in prices and shares in general equilibrium, we work out some analytical examples to provide more intuition. We also discuss how our results can be applied to economies with recursive dynamics.

Nested-CES economies. Household preferences are represented by a non-homothetic CES aggregator, which implies that budget shares vary according to (7). Recall that θ_0 is the elasticity of substitution across consumption goods and ε is the vector of income-elasticities. Production also uses nested-CES aggregators. Nested-CES economies can be written in many different equivalent ways, since they may have arbitrary patterns of nests. We adopt the following representation. We assume that each good $i \in N$ is produced with the production function

$$y_i = A_i G_i \left(\{m_{ij}\}_{j \in N}, \{l_{if}\}_{f \in F} \right) = A_i \left(\sum_{j \in N} \omega_{ij} m_{ij}^{\frac{\theta_i - 1}{\theta_i}} + \sum_{f \in F} \omega_{if} l_{if}^{\frac{\theta_i - 1}{\theta_i}} \right)^{\frac{\theta_i}{\theta_i - 1}},$$

where ω_{ij} and ω_{if} are constant parameters. Any nested-CES production network can be represented in this way if we treat each CES aggregator as a separate producer (see Baqaee and Farhi, 2019b).

Input-output matrix. We stack the expenditure shares of the representative household, all producers, and all factors into the $(1 + N + F) \times (1 + N + F)$ input-output matrix Ω. The first row corresponds to the household. To highlight the special role played by the representative agent, we index the household by 0, which means that the first row of Ω is equal to the household's budget shares introduced above ($\Omega_0 = b'$, with $b_i = 0$ for $i \notin N$).[28] The next N rows correspond to the expenditure shares of each producer on every other producer and factor. 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).

Leontief inverse matrix. The Leontief inverse matrix is the $(1 + N + F) \times (1 + N + F)$ matrix defined as

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

where Id is the identity matrix. The Leontief inverse matrix $\Psi \geq Id$ records the *direct and indirect* exposures through the supply chains in the production network. We partition Ψ in the following way:

$$\Psi = \left[\begin{array}{c|ccc|ccc} 1 & \lambda_1 & \cdots & \lambda_N & \lambda_{N+1} & \cdots & \lambda_{N+F} \\ \hline 0 & \Psi_{11} & \cdots & \Psi_{1N} & \Psi_{1N+1} & \cdots & \Psi_{1N+F} \\ 0 & & \ddots & & & & \\ 0 & \Psi_{N1} & & \Psi_{NN} & \Psi_{NN+1} & \cdots & \Psi_{NN+F} \\ \hline 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & 1 & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{array} \right].$$

[28] We expand the vector of demand-shifters $\Delta \log x$ and income elasticities ε to be $(1 + N + F) \times 1$, where $\Delta \log x_i = \varepsilon_i = 0$ if $i \notin N$.

The first row and column correspond to final demand (good 0). The first row is equal to the vector of sales shares for goods and factors λ'. The next N rows and columns correspond to goods, and the last F rows and columns correspond to the factors. Define the $(1 + N + F) \times F$ matrix Ψ^F as the submatrix consisting of the right F columns of Ψ, representing the network-adjusted factor intensities of each good. The sum of network-adjusted factor intensities for every good i is equal to one, $\sum_{f \in F} \Psi_{if} = 1$ because the factor content of every good is equal to one. In our results below we will use the identity that $\lambda' = b'\Psi$.

4.1 General characterization for nested-CES economies

According to Propositions 4 and 5, changes in real GDP and welfare can be computed by integrating observed and Hicksian sales shares with respect to technology shocks. For compact notation, we denote the welfare-relevant Hicksian variables implied by Definition 6 using the superscript *ev*. So, for example,

$$\lambda^{ev}(A, L) \equiv \lambda_i(A, L, u_{t_1}, x_{t_1}).$$

The following proposition pins down how observed and Hicksian sales shares, λ and λ^{ev}, vary as a function of the technology shocks. This proposition can then be used in combination with Propositions 4 and 5 to calculate exact changes in real GDP and welfare. For readability, we again assume away shocks to factor endowments and a path of taste shocks x and technology shocks A between t_0 and t_1.

Proposition 7 (Characterization for nested-CES economies)**.** *At any point in time t, changes in observed prices and the Leontief inverse are pinned down by the following equations:*

$$d \log p_{it} = - \underbrace{\sum_j \Psi_{ijt} d \log A_{jt}}_{\text{upstream TFP changes}} + \underbrace{\sum_{f \in F} \Psi_{ift} d \log \lambda_{ft}}_{\text{upstream factor price changes}} ,$$

$$d\Psi_{ilt} = \underbrace{\sum_j \Psi_{ijt}(\theta_j - 1) Cov_{\Omega_{(j,:),t}} \left(-d \log p_t, \Psi_{(:,l),t} \right)}_{\text{substitution effect}} + \mathbf{1}_{\{i=0\}} \underbrace{Cov_{\Omega_{(0,:),t}} \left(d \log x_t + \varepsilon_t d \log Y_t, \Psi_{(:,l),t} \right)}_{\text{taste shocks and income effect}}, \tag{17}$$

where changes in observed sales shares are given by $d\lambda_{it} = d\Psi_{0it}$ and changes in real GDP are given by $d \log Y_t = \sum_i \lambda_{it} d \log A_{it}$.

On the other hand, changes in welfare-relevant variables are pinned down by the following *system of differential equations*

$$d\log p^{ev}_{it} = -\sum_j \Psi^{ev}_{ijt} d\log A_{jt} + \sum_{f\in F} \Psi^{ev}_{ift} d\log \lambda^{ev}_{ft},$$

$$d\Psi^{ev}_{ilt} = \sum_j \Psi^{ev}_{ijt}(\theta_j - 1) Cov_{\Omega^{ev}_{(j,:),t}}\left(-d\log p^{ev}_t, \Psi^{ev}_{(:,l),t}\right), \tag{18}$$

where changes in welfare-relevant sales shares are given by $d\lambda^{ev}_{it} = d\Psi^{ev}_{0it}$ *and changes in welfare are given by* (14)*.*

For all of these expressions, the summations are evaluated over all goods and factors, so that i *and* $j \in \{0\} + N + F$, $Cov_{\Omega^{(j,:),t}}(\cdot)$ *is the covariance using the jth row of* Ω *at time* t *as the probability weights, and* $\Psi_{(:,i),t}$ *is the ith column of the Leontief inverse at time* t*.*

These differential equations can be solved by repeated iteration.[29] Once in possession of these paths, the change in real GDP and welfare are straightforward to calculate by cumulating the λ and λ^{ev}-weighted sum of technology shocks. For this iterative procedure, the boundary condition of the differential equations are that prices satisfy $p_{t_1} = p^{ev}_{t_1} = \mathbf{1}$ and the Leontief inverse matches $\Psi^{ev}_{t_1} = \Psi_{t_1}$.

For ex-post welfare questions, where the Leontief inverse Ψ is observed at t_1, we can calculate Ψ^{ev} between t_0 and t_1 by starting (18) at t_1 and going backwards to t_0. This process does not require knowledge of either the income elasticities ε nor the taste shocks $\Delta \log x$ since they do not appear in either the equation for $d\log p^{ev}$ nor the equation for $d\Psi^{ev}$.

Each term in the differential equations in Proposition 7 has a clear interpretation. We start by discussing the equation determining prices $d\log p$. This equation captures the fact that the price of each good $d\log p_i$ is determined by its (direct and indirect) exposure to the price of inputs j and factors f (captured by Ψ_{ijt} and Ψ_{ift} at time t).

On the other hand, the equation for $d\log \Psi_{ilt}$ shows that changes in the Leontief inverse are determined by substitutions by j, if j is an intermediary between i and l, as well as income and substitution effects if i is the household ($i = 0$). Finally, the welfare-relevant versions of these equations, $d\log p^{ev}$ and $d\Psi^{ev}$ are identical except that they take into account expenditure-switching due to income effects or taste shocks from the start, not along the transition.[30]

Remark 3 (Micro Welfare)**.** Proposition 7 can also be used to compute changes in microeconomic welfare EV^m. To do this, we compute the actual path of prices $d\log p$ using Proposi-

[29]When evaluating (17) between t_0 and t_1, we must take into account that income elasticities ε change with budget shares, as described in Appendix E.1.

[30]Proposition 7 generalizes Baqaee and Farhi (2019b) to economies with income effects and taste shocks.

tion 7 and then plug these price changes into (9). Unlike macroeconomic welfare EV^M, calculating $d \log p$ generically requires knowledge of both income elasticities and taste shocks.

To build more intuition, consider economies with only a single factor of production. In this case, the differential equations for $d \log p$ and $d \log p^{ev}$ are decoupled from the ones for $d\Psi$ and $d\Psi^{ev}$. This follows from the fact that the economy's single primary factor must have a sales share of unity. In other words, the following set of equations always hold when the economy has a single factor: $\lambda_f = \lambda_f^{ev} = \Psi_{0f} = \Psi_{0f}^{ev} = 1$ for $f \in F$. This allows for a simple characterization of both welfare and real GDP up to a second-order approximation.

Proposition 8 (Approximate Macro Welfare vs GDP: Single Factor)**.** *Consider some perturbation in technology, $\Delta \log A$, and tastes, $\Delta \log x$. When the economy has one factor of production, the change in real GDP is*

$$\Delta \log Y \approx \sum_{i \in N} \lambda_i \Delta \log A_i + \frac{1}{2} \sum_j \lambda_j (\theta_j - 1) Var_{\Omega_{(j,:)}} \left(\sum_{i \in N} \Psi_{(:,i)} \Delta \log A_i \right)$$
$$+ \frac{1}{2} Cov_{\Omega_{(0,:)}} \left(\Delta \log x + (\sum_{i \in N} \lambda_i \Delta \log A_i) \varepsilon, \sum_{i \in N} \Psi_{(:,i)} \Delta \log A_i \right), \tag{19}$$

and the difference between welfare and GDP is

$$EV^M - \Delta \log Y \approx \frac{1}{2} Cov_{\Omega_{(0,:)}} \left(\Delta \log x + (\sum_{i \in N} \lambda_i \Delta \log A_i) \varepsilon, \sum_{i \in N} \Psi_{(:,i)} \Delta \log A_i \right), \tag{20}$$

where λ, Ω, Ψ, and ε are evaluated at t_0.

Proposition 8 is a general equilibrium counterpart to Proposition 2. We discuss (19) and (20), starting with (19). The first term in Equation (19) is the Hulten-Domar term. The other terms are second-order terms resulting from the fact that sales shares change in response to shocks. The first one of these terms captures nonlinearities due to the fact that sales shares can respond to changes in relative prices caused by technology shocks (these effects are emphasized by Baqaee and Farhi, 2019b). The terms on the second line of (19), which are the ones we focus on in this paper, capture changes in sales shares due to changes in tastes or non-homotheticities.

Equation (20) shows that while real GDP correctly accounts for substitution due to supply shocks, in order to measure welfare, it needs to be corrected for expenditure-switching due to changes in final demand caused by taste shocks or income effects. Whereas in partial equilibrium, the gap between welfare and real GDP is proportional to the covariance of supply and demand shocks (see Proposition 2), equation (20) shows that in general equilibrium, the relevant statistic is the covariance of demand shocks with a network-adjusted notion of supply shocks, and not supply shocks per-se.[31]

4.2 Analytical Examples

We now work through some simple examples to illustrate the forces that drive a gap between λ and λ^{ev} and, by extension, real GDP and welfare.

Example 1 (Correlated Supply and Demand Shocks)**.** We start with the simplest possible example, a one sector model without any intermediates. In this case, sales shares are just budget shares $\lambda_i = b_i = \Omega_{0i}$. Therefore, Proposition 8 simplifies to

$$EV^M - \Delta \log Y \approx \frac{1}{2} \left(Cov_b \left(\Delta \log x, \Delta \log A \right) + Cov_b \left(\varepsilon, \Delta \log A \right) \mathbb{E}_b[\Delta \log A] \right). \tag{21}$$

Welfare changes are greater than the change in real GDP if productivity and demand shocks (i.e. shifts in demand curves) are positively correlated. This could happen either because preferences exogenously change to favor high productivity growth goods, $Cov_b \left(\Delta \log x, \Delta \log A \right) > 0$, or income effects favor high productivity growth goods, $Cov_b \left(\varepsilon, \Delta \log A \right) \Delta \log Y > 0$. When shifts in demand are orthogonal to shifts in supply, to a second-order approximation, real GDP measures welfare correctly. By Proposition 3, in this example, macro welfare EV^M is the same as micro welfare EV^m.

We now work through some simple examples with multiple factors of production to illustrate how nonlinear PPFs affect the previous results.

Example 2 (Decreasing Returns to Scale)**.** Consider the one-sector model without intermediate inputs in Example 1 but now suppose that production functions are non-constant-returns-to-scale. Specifically, the production for good i is

$$y_i = A_i L_i^{\gamma},$$

where L_i is labor and γ need not equal 1. To apply our propositions to this economy, where producers have non-constant-returns production functions, we introduce a set of producer-specific factors in inelastic supply, and suppose that each producer has a Cobb-Douglas production function that combines a common factor with elasticity γ and a producer-specific factor with elasticity $1 - \gamma$. This means that our economy has $1 + N$ factors.

[31] Equation (20) shows that the gap between welfare and real GDP does not depend on elasticities of substitution. This is a consequence of the fact that the PPF in Proposition 8 is linear and expenditure-switching does not affect relative prices.

For simplicity, suppose that preferences are homothetic ($\varepsilon_i = 1$ for every i), but potentially unstable ($\Delta \log x \neq 0$). We apply Proposition 6 to compute the difference between welfare and real GDP. To do this, we first use Proposition 7 to compute changes in observed and Hicksian sales shares due to demand shocks and then plug this into Proposition 6 to get the difference between welfare and real GDP up to a second order approximation:

$$EV^M - \Delta \log Y \approx \frac{1}{2} \frac{Cov_b(\Delta \log x, \Delta \log A)}{1 + (\theta_0 - 1)(1 - \gamma)}. \tag{22}$$

This expression simplifies when we have constant-returns to scale ($\gamma = 1$) or when the elasticity of substitution is unity ($\theta_0 = 1$). In these cases, following Proposition 3, the PPF is linear. Outside of these cases, complementarities ($\theta_0 < 1$) amplify the impact of preference shocks under decreasing returns to scale ($\gamma < 1$). Intuitively, if preferences shift in favor of some good, the price of that good rises due to decreasing returns to scale. The fact that the price of the good increases raises the sales share of that good due to complementarities, which creates a feedback loop, raising prices of the good further, and causing additional substitution. In other words, in the decreasing returns to scale model with complementarities, sales shares respond more strongly to demand shocks. Given that sales shares respond more strongly to demand shocks, the necessary adjustment to correct real GDP is larger.

Example 3 (Macro vs. Micro Welfare Change)**.** Finally, we demonstrate the difference between macro and micro welfare changes using the previous example. The economy in the previous example has multiple factors and unstable preferences. Therefore, macro and micro notions of welfare are different since the PPF is no longer linear.

To illustrate this difference, suppose that only preference shocks are active (there are no supply shocks $\Delta \log A = 0$ and $\Delta \log L = 0$). Since the PPF is being held constant, macro-welfare changes are also zero. Micro-welfare changes, on the other hand, are not equal to zero. Specifically, by Proposition 2, micro welfare improves $EV^m > 0$ if preference shocks negatively covary with price changes. Using Proposition 7, changes in prices are

$$d \log p_i = \frac{(1 - \gamma)}{(1 + (\theta_0 - 1)(1 - \gamma))} \left(d \log x_i - \mathbb{E}_b[d \log x] \right).$$

If there are decreasing returns, $\gamma < 1$, then a positive demand shock for i raises the price of i. This initial change in the price is amplified by general equilibrium forces if goods are complements and mitigated if goods are substitutes (this is the multiplier in the denominator). We can now apply Proposition 2 to obtain micro welfare, up to a second order,

$$EV^m \approx -\frac{1}{2}\frac{(1-\gamma)}{(1+(\theta_0-1)(1-\gamma))}Var_b(\Delta\log x) \neq 0 = EV^M.$$

With decreasing returns to scale ($\gamma < 1$), micro welfare is negative since the demand shock increases the prices of goods the consumer now values more. From a micro perspective, where the agent takes the budget sets as given, the agent is worse off. Of course, from a societal perspective, welfare has not changed, since the production possibility set of the economy has not changed.

Appendix G contains additional examples showing how input-output connections can amplify or mitigate the gap between macro welfare EV^M and real GDP $\Delta \log Y$.

4.3 Dynamic Economies

As mentioned earlier, at an abstract level, all of our results can be applied to dynamic economies by using the Arrow-Debreu formalism. In particular, we can index goods by period of time and state of nature and apply our results to these economies (see e.g. Basu et al., 2012). In a dynamic economy the utility function is intertemporal and capital accumulation must be treated as an intertemporal intermediate good, as advocated by Barro (2021). Proposition 5 implies that, in such a model, macro welfare can be computed using the final (intertemporal) indifference curve of the representative agent.

In this section, we specialize these results further to show how Proposition 5 can be used to make steady-state to steady-state welfare comparisons in dynamic models with unstable and non-homothetic preferences. Specifically, we consider a dynamic multi-sector model with production of consumption goods and investment goods similar to the models that are often used to study structural transformation (Herrendorf et al., 2013). For simplicity, we abstract from growth and restrict our discussion to non-homothetic CES preferences.

Consider a perfectly competitive dynamic economy indexed by the initial period t with a representative agent whose intertemporal preferences are given by

$$\mathcal{U}_t = \sum_{s=t}^{\infty}\beta^{s-t}u(C_s), \qquad \sum_i \omega_{i0}x_{it}\left(\frac{c_{is}}{C_s^{\xi_i}}\right)^{\frac{\theta_0-1}{\theta_0}} = 1,$$

where C_s is a non-homothetic (and potentially unstable) CES aggregator. The economy has the same set of goods every period, and every good i in period s is produced according to constant returns production technology with arbitrary input-output connections

$$y_{is} = A_{is} G_i \left(\{m_{ijs}\}_{j \in N}, H(l_{is}, k_{is}) \right),$$

where A_{is} is a productivity shifter, l_{is} and k_{is} are labor and capital inputs, and H is constant returns to scale.

Labor L_s in each period is inelastically supplied, and capital is accumulated according to a capital accumulation technology

$$K_{s+1} = (1 - \delta)(K_s + X_s),$$

where X_s is aggregate investment. Investment goods are produced according to a constant returns technology with arbitrary input-output connections

$$X_s = A_{Is} X \left(\{m_{Ijs}\}_{j \in N}, H(l_{Is}, k_{Is}) \right).$$

The intertemporal PPF of economy t is defined by an initial capital stock inherited from the past, a path of future labor endowments, and a path of vectors of productivities: $(K_t, \{L_s\}_{s=t}^{\infty}, \{A_s\}_{s=t}^{\infty})$. This economy has infinitely many factors: the initial capital stock and the path of labor endowments $(K_t, \{L_s\}_{s=t}^{\infty})$. The welfare change between t_0 and t_1 is the proportional change in factor endowments of the t_0 economy required to make the household indifferent between that and the t_1 economy. We say that economy t is in *steady-state* if the vector of productivities A_s, labor endowments L_s, per-period utility $u(C_s)$, and capital stocks K_s are constant over time.

The following proposition shows that computing the welfare change between t_0 and t_1 is straightforward if the economy is in steady-state in both t_0 and t_1.

Proposition 9 (Dynamic Welfare Change)**.** *Consider two dynamic economies, denoted t_0 and t_1, that are in steady-state. The change in macro welfare is given by*

$$EV^M = \log \left(\frac{\sum_i p_{it_1} c_{it_1}}{\sum_i p_{it_0} c_{it_0}} \right) + \log \left(\sum_i b_{it_1} \left(\frac{p_{it_0}}{p_{it_1}} \right)^{1-\theta_0} \right)^{\frac{1}{1-\theta_0}}. \tag{23}$$

In words, macroeconomic welfare in this dynamic economy is equal to the change in nominal consumption expenditures deflated by the exact-algebra CES price index associated with the t_1 indifference curve, exactly as for the partial equilibrium microeconomic welfare in expression (9), despite the fact that this is a dynamic general equilibrium economy with infinitely many factors.

5 Applications

In this section we consider three applications of our results. The first application is about long-run growth, quantifying the difference between welfare-relevant and measured aggregate productivity growth in the presence of income effects and demand instability. The second application is about short-run fluctuations, showing that correlated product-level supply and demand shocks within industries drive a wedge between measured real consumption and welfare even in the short-run. Our final application is a business cycle event study, where we use the Covid-19 recession to demonstrate the difference between macroeconomic and microeconomic welfare and how demand instability can make measured real GDP an unreliable metric for changes in production.

5.1 Application I: Long-Run Growth and Structural Transformation

As economies grow, sectors with low productivity growth tend to expand compared to sectors with faster productivity growth. This means that over time, aggregate productivity growth is increasingly determined by those sectors whose productivity growth is slowest. This phenomenon is oftentimes called Baumol's cost disease.

Following Nordhaus et al. (2008), aggregate productivity growth between t_0 and t_1 can be decomposed into two terms:

$$\Delta \log TFP = \sum_{t=t_0}^{t_1} \sum_{i \in N} \lambda_{it_0} \Delta \log A_{it} + \underbrace{\sum_{t=t_0}^{t_1} \sum_{i \in N} \left(\lambda_{it} - \lambda_{it_0}\right) \Delta \log A_{it}}_{\text{Baumol Adjustment}},$$

where λ_{it} is the sales shares of industry i in period t and $\Delta \log A_{it}$ is the growth in gross-output productivity over each time period.[32] The first term captures changes in aggregate TFP if industry-structure had remained fixed, and the second term is the adjustment attributed to the fact that sales shares change over time. The second-term captures the importance of Baumol's cost disease.[33]

[32]Technically, this is an approximation, since we define aggregate TFP in continuous time but the data is measured in discrete time (at annual frequency). However, this approximation error, resulting from the fact that the Riemann sum is not exactly equal to the integral is likely to be negligible in practice. At our level of disaggregation, long run TFP growth is very similar if we weight sectors using sales shares at time t or time t and $t+1$ averages.

[33]For this exercise, we abstract from investment decisions and apply our formulas statically. This means that we assume a reduced-form representation whereby preference relations are defined over all final goods in a given period (including government spending, net exports, and investment) and calculate welfare changes between two time periods taking preferences, technologies, and factor quantities as given. We also calculate welfare using consumption data in Appendix H.3, using Proposition 9.

Proposition 5 implies that, for the purposes of welfare, changes in sales shares due to income effects or demand instability must be treated differently to changes in sales shares due to substitution effects. In particular, for EV^M, the relevant measure of the change in TFP is

$$\Delta \log TFP^{ev} = \sum_{t=t_0}^{t_1} \sum_{i \in N} \lambda_{it_0} \Delta \log A_{it} + \underbrace{\sum_{t=t_0}^{t_1} \sum_{i \in N} \left(\lambda_{it} - \lambda_{it_0}\right) \Delta \log A_{it}}_{\text{Baumol Adjustment}} + \underbrace{\sum_{t=t_0}^{t_1} \sum_{i \in N} \left(\lambda_{it}^{ev} - \lambda_{it}\right) \Delta \log A_{it}}_{\text{Welfare Adjustment}},$$

where λ^{ev} is the Hicksian sales-shares of each industry holding fixed final preferences and income-level.

Two polar extremes. Computing the welfare adjustment term to obtain $\Delta \log TFP^{w}$ requires an explicit structural model of the economy. However, there are two polar cases in which $\Delta \log TFP^{w}$ can be calculated without specifying the detailed model. The first extreme is when demand is stable and homothetic, and changes in sales shares are due only to relative price changes (substitution effects). The second extreme is when there are no substitution effects in sales shares (as in a Cobb-Douglas economy), and changes in sales shares are only due to income effects or demand instability. If structural transformation is driven by a combination of substitution effects and non-homotheticities or demand instability, then the change in welfare TFP will be somewhere in between these two cases, as discussed in Appendix H. The following corollary of Proposition 5 summarizes the change in welfare-TFP in these two polar cases.

Corollary 1. *If changes in sales shares are due only due only to substitution effects, then*

$$\Delta \log TFP^{ev} = \Delta \log TFP = \sum_{t=t_0}^{t_1} \sum_{i \in N} \lambda_{it} \Delta \log A_{it}.$$

If changes in sales shares are due only to non-homotheticity or instability of demand, then

$$\Delta \log TFP^{ev} = \Delta \log TFP + \sum_{t=t_0}^{t_1} \sum_{i \in N} \left(\lambda_{it_1} - \lambda_{it}\right) \Delta \log A_{it} = \sum_{t=t_0}^{t_1} \sum_{i \in N} \lambda_{it_1} \Delta \log A_{it}.$$

In the first case, since preferences are homothetic and stable, welfare-TFP is equal to TFP in the data. In the second case, since there are no substitution effects in production or demand, compensated sales shares do not respond to productivity changes, so $\lambda_{it}^{ev} = \lambda_{it_1}$.

To quantify Corollary 1, we use US-KLEMS data on sales shares and TFP growth for

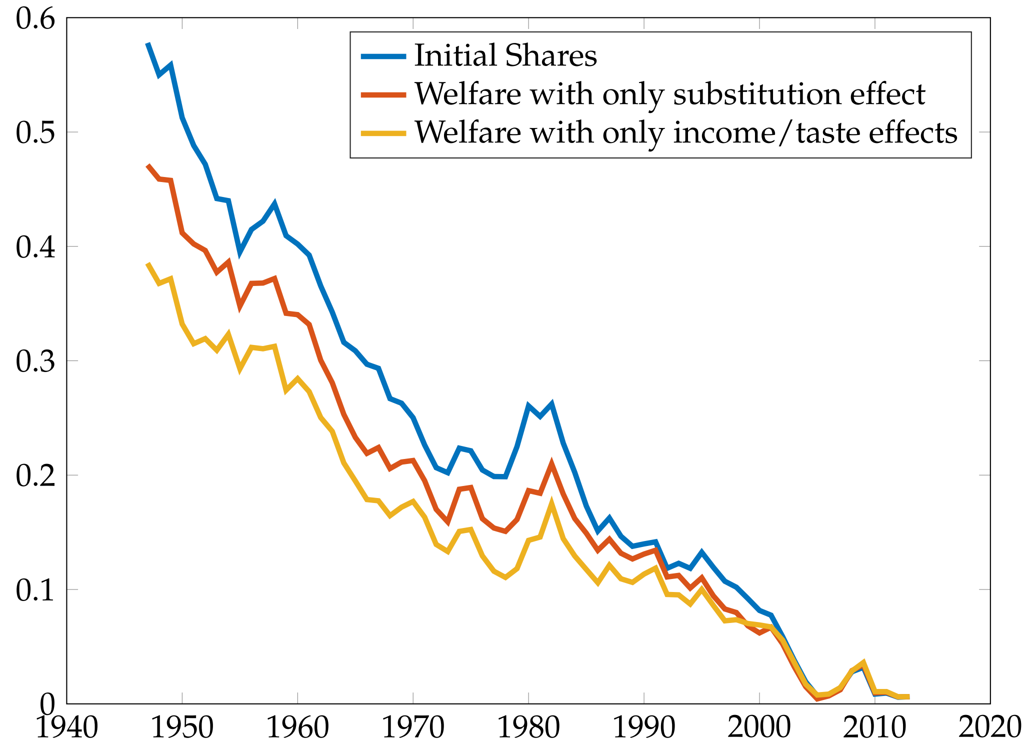


Figure 2: Growth in welfare-relevant TFP (in logs) from 1947 to 2014 using US-KLEMS. The blue line uses initial shares (in each year t between 1947 and 2014) to calculate TFP changes. The red and yellow line measure the increase in welfare-relevant TFP between t and 2014 under alternative assumptions about income and substitution elasticities. The red line assumes that sales shares change only due to substitution effects (welfare-relevant TFP is equal to measured chained-aggregate TFP). The yellow line assumes that sales shares change only due to income effects (or demand instability).

61 private-sector industries between 1947 and 2014. We calculate changes in industry-level gross-output TFP following the methodology of Jorgenson et al. (2005) and Carvalho and Gabaix (2013).[34] Figure 2 plots EV^M comparing 2014 to previous years under alternative assumptions about substitution and income elasticities. For comparisons that are relatively close to 2014, the change in welfare is not very sensitive to our assumptions about elasticities. This is because at high frequency, the shocks are small and the sales shares are reasonably stable. However, the assumptions about substitution and income elasticities do start to play a role as we roll the comparison back farther in time. Comparing 1947 to 2014, the constant-initial-sales-share term grows by around 58 log points (or 78%), whereas the chain-linked change in aggregate TFP grew by around 47 log points (or 60%). Hence, Baumol's cost-disease caused aggregate TFP to fall by 10 log points, and reduced aggregate productivity growth by around 23 percent (from 78% to 60%).

If we assume that structural transformation is due solely to income effects and taste

[34]For each industry, the change in TFP is itself a chain-weighted index calculated as output growth minus share-weighted input growth. Inputs are industry-level measures of materials, labor, and capital services.

shocks, then by Corollary 1 the growth in welfare-relevant TFP from 1947-2014 was 37 log points (or 46%) instead of the measured 47 log points (or 60%) — that is, to say, a 23 percent additional reduction in the growth rate.

Intuitively, welfare-based productivity increases less than TFP because, relative to 1947, preferences in 2014 favor low productivity growth sectors such as services (due to either income effects or demand instability). This means that, at 1947 prices, households require less income growth to be indifferent between their budget constraint in 1947 and the one in 2014. This is because sectors with high income elasticities or that consumers prefer in 2014, like services, were cheaper compared to manufacturing in 1947 than in 2014.[35]

To sum up, structural transformation caused by income effects or demand instability reduced welfare, EV^M, by roughly twice as much as structural transformation caused by substitution effects. The fact that income effects or demand instability are twice as important as substitution effects is not a coincidence and can be understood using the second-order approximation in Proposition 6. The former are multiplied by 1 whereas the latter are multiplied by 1/2. Since for this application, changes in Domar weights have roughly linear time-trends, a second-order approximation performs well.

We elaborate on this point further in Appendix H and also provide some quantitative illustrations away from the two polar extremes we discussed above. In this appendix, we compute welfare changes for different values of elasticities of substitution in consumption and production using Proposition 7. We show that welfare-relevant TFP is closer to measured TFP if the elasticity of substitution across heavily disaggregated industries (in consumption or production) is significantly lower than one.[36]

5.2 Application II: Aggregation Bias with Product-Level Taste Shocks

In the previous application, we considered a long-run industry-level application. With industry-level data, the gaps between welfare and chain-weighted indices are usually modest over the short-run because industry-level sales shares are relatively stable at high frequency (e.g. see Figure 2). However, this does not mean that these biases are necessarily

[35]This intuition is flipped for compensating variation at 1947 preferences. Preferences in 1947 favor manufacturing over services. Therefore, at 2014 prices, households require a larger reduction in 2014 income to make them indifferent to 1947. More generally, if structural transformation is purely due to income effects or preference instability, then welfare-based productivity growth using CV at initial preferences is given by initial sale-share weighted productivity growth, $\sum_{t=t_0}^{t_1} \sum_{i \in N} \lambda_{it_0} \Delta \log A_{it}$ (which corresponds to the Initial Shares line in Figure 2), so for this exercise, the Baumol adjustment is not welfare-relevant. The fact that EV at final preferences and CV at initial preferences are different stems from the fact that they answer different questions, so EV uses demand from 2014 whereas CV uses demand from 1947. These demand curves differ if preferences are unstable/non-homothetic.

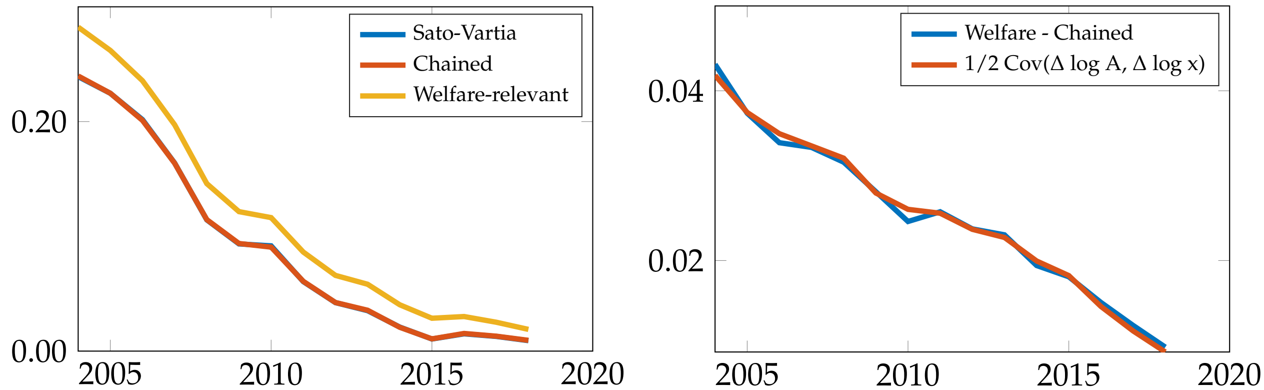
[36]Appendix H.3 shows that similar conclusions apply if we measure welfare changes using BEA data on Personal Consumption Expenditures prices and budget shares across goods in the US.

absent from short-run data.

Whereas industry sales shares are stable at high frequency, firm or product-level sales shares are highly volatile even over the very short-run. If firms' or products' supply and demand shocks are correlated, then measured industry-level output is biased relative to what is relevant for welfare. In Appendix I, we formally show that the biases in industry-level data are not diversified away as we aggregate, even if all products are infinitesimal in their industry and all industries are infinitesimal in the aggregate economy. Furthermore, we provide conditions under which the within-industry biases are, to a second-order, linearly separable from the across-industry biases. That is, the overall bias is the sum of the cross-industry bias (that we studied in the previous section) plus additional biases driven by within-industry covariance of supply and demand shocks. If supply and demand shocks at the product level are persistent, then the covariance of supply and demand shocks, and hence the bias, is larger over longer horizons.

We provide an empirical illustration of the magnitude of the biases caused by taste shocks in product-level data using the Nielsen Consumer Panel database. In the body of the paper, we only briefly describe the dataset, and refer readers to Appendix J for more details. The Nielsen Consumer Panel tracks the purchasing behavior of about $40,000$ to $60,000$ panelists every year from 2004 to 2019 as they shop in a wide variety of non-durable consumer goods (food, non-food groceries, general merchandise, etc.). A product in the data is defined by its unique Universal Product Code (UPC), and each product is assigned to a module. Our balanced sample covers roughly 820 modules. Panelists in the sample are assigned weights, allowing purchases by the panel to be projected to a nationally representative sample.

We model national demand for UPCs in a given module using a homothetic but unstable CES functional form. We set $t_1 = 2019$ and then for each $t_0 < 2019$, we calculate a welfare-relevant deflator by module for continuing goods using preferences in 2019 following equation (9). The price of each UPC in each year is calculated as the ratio of national expenditures on that UPC over units sold over the whole year. For each t_0, we include only UPCs purchased in each quarter of each year between t_0 and t_1. In other words, we abstract from product entry and exit by focusing on the continuing-goods price index (see Section 6 for how to deal with product entry-exit when preferences are unstable). For the same set of UPCs, we also compute the change in inflation as measured by a chained Tornqvist index (a discrete time approximation to the Divisia index) as well as the commonly used Sato-Vartia (SV) index. We then combine these module-level inflation rates into a single number by weighing each module according to its share in overall expenditures in the year 2019, which corresponds to assuming demand across modules is Cobb-Douglas.



(a) Inflation rates

(b) Bias

Figure 3: Welfare-relevant, chain-weighted, and Sato-Vartia inflation rate for continuing products. The chain-weighted and Sato-Vartia index are almost the same. The welfare-relevant rate is computed assuming that the elasticity of substitution across UPCs in the same module is 4.5.

Figure 3 displays the results assuming that the elasticity of substitution across UPCs in the same module is 4.5.[37] Panel 3a shows the welfare-relevant, chained, and SV inflation rates for each t_0. The chained and SV index move closely with each other, but the welfare-relevant inflation rate is higher. Starting in 2018, welfare-relevant inflation is around 1 percentage point higher than the chained and SV indices. Intuitively, this is because changes in prices and changes in demand residuals are positively correlated, and hence, following the logic of Proposition 2, the chained index understates inflation. The gap increases as we go back further in time and over the whole sample, the gap widens to 4.3 percentage points.

Panel 3b plots the difference between the welfare-relevant and chained indices and the expected bias term from Proposition 8 (or Proposition 13 in Appendix I which explicitly models the industry-firm structure). To calculate UPC-level productivity shocks $\Delta \log A$ we use the change in each UPC's log price relative to the average log price change for that module. We construct Δx as the difference between the observed change in expenditure shares and the change in expenditure share implied by a CES aggregator with elasticity $\theta_0 = 4.5$.

Panel 3b shows that the second-order approximation performs well, as the gap between the welfare-relevant and chained index is approximately equal to half the covariance between supply and demand shifters. Furthermore, the gap widens as we extend the time

[37] Estimating the elasticity of substitution is beyond the scope of this paper, therefore, for our empirical illustration we draw on estimates from the literature. An elasticity of 4.5 is at the lower range of estimates reported by Redding and Weinstein (2020) and Jaravel (2019). In Appendix J, we report results for higher and lower elasticities. We find that the size of the bias is increasing in the elasticity of substitution. In this sense, the results in Figure 3 are relatively conservative.

horizon because the covariance of supply and demand shifters increases. This is natural if supply and demand shocks are persistent. Hence, these disaggregated product-level biases, which are assumed away if one begins with more aggregated data, are non-negligible.

We report robustness with respect to the elasticity of substitution parameter in Appendix J. In this appendix, we show that the size of the bias gets smaller as we get closer to Cobb-Douglas. This is because in the data changes in prices and changes in expenditure shares are approximately uncorrelated. When demand is Cobb-Douglas, changes in expenditure shares are driven only by taste shocks, and so taste shocks are roughly uncorrelated with price changes. Hence, following the logic of Proposition 2, the bias is smaller in the Cobb-Douglas case and larger if the calibrated elasticity is greater than 4.5.

5.3 Application III: the Covid-19 Recession

Our final application examines how real GDP, microeconomic welfare, and macroeconomic welfare were affected during the Covid-19 recession. The Covid-19 recession is an interesting case study since sectoral expenditure shares changed substantially during this time, these changes were not explainable via changes in observed prices alone, and the movements in demand curves were correlated with movements in supply curves. These are exactly the conditions under which micro welfare, macro welfare, and real GDP can diverge from each other.

Cavallo (2020) argues that, during this episode, the fact that price indices were not being chained at high enough frequency led to "biases" in official measures of inflation. However, since final demand was unstable during this period, chaining is not theoretically justified. As we have argued, chaining is only theoretically valid if expenditure-switching is caused by substitution effects, and not if expenditure-switching is caused by shocks to demand. Furthermore, if changes in prices are themselves caused by changes in demand (due to decreasing returns to scale), then microeconomic welfare and macroeconomic welfare changes are different.

In this section, we do not attempt to measure the welfare costs of Covid-19 itself. This is because households do not make choices over whether or not they live in a world with Covid-19. Therefore, their preferences about Covid-19 itself are not revealed by their choices. Instead, we ask a more modest question: how does the household value changes in prices (micro welfare) and changes in production (macro welfare), holding *fixed* the presence of Covid-19.

To study this episode, we use a modified version of the quantitative model introduced in Section 4. Since we are interested in a short-run application, we assume that factor markets are segmented by industry, so that labor and capital in each industry is inelastically supplied. We calibrate share parameters to match the 71 industry US input-output table in 2018 (we exclude government sectors) from the BEA, and consider a range of elasticities of substitution. Following Baqaee and Farhi (2020), we model the Covid-19 recession as a combination of negative sectoral employment shocks and sectoral taste shifters. We hit the economy with a vector of primitive supply and demand shocks. The reductions in sectoral employment are calibrated to match peak-to-trough reductions in hours worked by sector from January, 2020 to May, 2020. The primitive demand shifters are calibrated to match the observed peak-to-trough change in personal consumption expenditures by sector from January, 2020 to May, 2020 (conditional on the supply shocks and the elasticities of substitution).[38]

We consider three different calibrations informed by empirical estimates from Atalay (2017) and Boehm et al. (2015): high complementarities, medium complementarities, and no complementarities (Cobb-Douglas). The high complementarity scenario sets the elasticity of substitution across consumption goods to be 0.7, the one across intermediates to be 0.01, across value-added and materials to be 0.3, and the one between labor and capital to be 0.2. The medium complementarities case sets the elasticity of substitution across consumption goods to be 0.95, the one across intermediates to be 0.01, across value-added and materials to be 0.5, and the one between labor and capital to be 0.5. The Cobb-Douglas calibration sets all elasticities of substitution equal to unity.

Table 1 displays welfare changes between January 2020 and May 2020 in the calibrated model. We report separately micro and macro welfare based on pre-Covid (initial, Q1-2018) and post-Covid (final, Q2-2020) preferences. Recall that micro and macro welfare are not equal in this economy because the PPF is nonlinear.

Table 1 shows that the drop in micro welfare is larger under post-Covid preferences than under pre-Covid preferences. This is because, as shown in Example 3, demand shocks reduce micro welfare in the presence of decreasing returns to scale. Intuitively, if there are decreasing returns to scale, then demand shocks increase the price of goods that consumers value more and this causes micro welfare to drop (since whatever households value becomes more expensive relative to the past).

This pattern is exactly reversed for macro welfare. Macro welfare is higher at post-Covid preferences than at pre-Covid preferences. This is because the negative supply shocks were biggest in those sectors where demand also fell more drastically (e.g. transportation and energy). Hence, the reduction in welfare is smaller with post-Covid pref-

[38]Changes in labor by sector and personal consumption expenditures, used to calibrate supply and demand shocks, are taken from Baqaee and Farhi (2020). For related analysis of Covid-19 induced supply shocks, see e.g. Bonadio et al. (2020) and Barrot et al. (2020). For related analysis of Covid-19 induced demand shocks, see Cakmakli et al. (2020).

Table 1: The change in micro and macro welfare with pre-Covid and post-Covid preferences given the supply and demand shocks between February 2020 to May 2020. Chained real consumption is computed assuming supply and demand shocks arrive simultaneously.

Elasticities	High compl.	Medium compl.	Cobb-Douglas
Micro pre-Covid preferences	-11.7%	-9.1%	-8.7%
Micro post-Covid preferences	-13.2%	-12.3%	-10.9%
Macro pre-Covid preferences	-16.2%	-12.5%	-10.8%
Macro post-Covid preferences	-10.1%	-9.4%	-9.0%
Chained real consumption	-12.1%	-10.6%	-9.8%

erences because those goods that the economy is less capable of producing are less desirable. This illustrates that micro and macro welfare answer different questions, and the answers to these questions can be quantitatively very different. Furthermore, comparing columns of Table 1 shows that the magnitude of these differences depend on the details of the production structure like the extent of complementarities in production. As we raise the elasticities of substitution in production closer to unity (Cobb-Douglas), the differences between macro and micro notions become less dramatic. This is because the PPF becomes less curved.

In Table 1, we also compute real consumption assuming supply and demand shocks arrive simultaneously and linearly over time. Interestingly, chained real consumption in Table 1 does not exactly measure any of the different welfare notions. This is because supply and demand shocks are not orthogonal along the path. In fact, if we change the order or path of supply and demand shocks, real consumption changes value (even though the initial and final allocation are not changing). For example, if the supply shocks arrive before the demand shocks, then real consumption equals macro welfare changes at pre-Covid preferences. On the other hand, if demand shocks arrive before the supply shocks, then real consumption equals macro welfare changes at post-Covid preferences.[39]

Hence, if the supply and demand shocks do not disappear in exactly the same way as they arrived, measured real consumption (or GDP) after the recovery can be higher or lower than it was before the crisis, even if the economy returns exactly to its pre-Covid allocation. If in the downturn, demand shocks arrive before supply shocks (so real consumption falls by roughly 10% in the high complementarities case, according to Table 1) and, in the recovery, demand shocks disappear before the supply shocks (so real consump-

[39]These two observations follow from the fact that demand shocks on their own have no impact on real GDP (see Proposition 4) and, conditional on fixed and homothetic preferences, real GDP equals macro welfare.

tion rises by roughly 16%), then real consumption is as much as 6% higher when comparing pre-shock real consumption to post-recovery real consumption. This is despite the fact that every price and quantity is the same when comparing the pre-shock allocation to the post-recovery allocation. Hence, during episodes where final demand is unstable, chained real GDP and consumption are unreliable guides for measuring output or welfare, even if we chain in continuous time.[40]

This example highlights that when comparing chain-weighted measures in the data against chain-weighted measures constructed using model-generated data, the specific path of shocks between two periods must be taken into account and matters if the model features non-homotheticities and/or taste shifters.

6 Extensions

In this section, we briefly summarize how our theoretical results can be extended in to account for new goods, non-CES functional forms, heterogeneous agents, and inefficiencies.

Extensive margin.

In our discussion, we do not explicitly deal with the new goods problem (i.e. goods discontinuously appearing or disappearing). As Fisher and Shell (1968) point out, the classic treatment of new goods is somewhat uncomfortable: "*The consumer is assumed to have always had an unchanging preference relation, complete with axes for all new goods of whose potential existence he in fact was not aware before their introduction.*" One might think that, at least in some cases, consumers consume a new good because of a change in tastes rather than a change in the price. In this section, using the commonly used CES specification, we show that the welfare implications of this can be profound.

Consider a consumer whose preferences are homothetic CES with taste shifters x and elasticity of substitution $\theta_0 > 1$. Good i is *unavailable* if its price is infinite, $p_i = \infty$, and *available* otherwise. Demand for good i may be zero either because the good is unavailable ($p_i = \infty$) or because the consumer does not value the good ($x_i = 0$).

Split the set of goods that consumers value at t_1 and are available in either t_0 or t_1 into

[40] This is related to a problem known as *"chain drift"* bias in national accounting. Chain drift occurs when a chained index registers an overall change between t_0 and t_1 even though all prices and quantities in t_0 and t_1 are identical. This is a specific manifestation of path dependence of chained indices (see Hulten,1973) and, by the gradient theorem for line integrals, it must be driven by either demand instability, income effects, or approximation errors due to discreteness. Chain drift bias can thus appear when movements in prices and quantities are oscillatory, where changes that take place over some periods are reversed in subsequent periods. Welfare changes do not exhibit chain drift since, by definition, they depend only on t_0 and t_1 variables.

three sets: (1) $\mathcal{C}$: continuing goods consumed in both periods, $b_{it_0} > 0$ and $b_{it_1} > 0$; (2) $\mathcal{N}$: newly consumed goods that either were unavailable at t_1 ($p_{it_0} = \infty$ and $p_{it_1} < \infty$) or that were available ($p_{it_0} < \infty$) but are valued at t_1 and not at t_0 ($x_{it_0} = 0$ and $x_{it_1} > 0$); (3) $\mathcal{X}$: exiting goods that become unavailable at t_1, $p_{it_0} < \infty$ and $p_{it_1} = \infty$.[41]

The following proposition derives the change in welfare.

Proposition 10 (Welfare with Product Entry/Exit)**.** *The change in equivalent variation at t_1 tastes is given by*

$$EV^m = \log \frac{I_{t_1}}{I_{t_0}} - \log \frac{P_{t_1}}{P_{t_0}}, \tag{24}$$

where

$$\frac{P_{t_1}}{P_{t_0}} = \left(\frac{b^c_{t_1} \left(\frac{P^c_{t_1}}{P^c_{t_0}} \right)^{\theta_0 - 1} + \left(1 - b^c_{t_1} \right) \left(\frac{P^n_{t_1}}{P^n_{t_0}} \right)^{\theta_0 - 1}}{1 - b^{x_{t_1}}_{t_0}} \right)^{\frac{1}{\theta_0 - 1}}. \tag{25}$$

In this expression $b^c_{t_1}$ is the expenditure share on continuing goods in t_1, and $P^c_{t_1}/P^c_{t_0}$ and $P^n_{t_1}/P^n_{t_0}$ is the change in a CES price index for continuing, $\mathcal{C}$, and new, $\mathcal{N}$, goods under t_1 tastes. Finally, $b^{x_{t_1}}_{t_0}$ is the expenditure share on exiting goods under t_1 tastes and p_{t_0} prices.

Applying (25) requires three pieces of information:

i. The t_1 share of continuing goods, $b^c_{t_1}$, and changes in the t_1 price index for continuing goods $P^c_{t_1}/P^c_{t_0}$.

ii. The price index for newly consumed goods $P^n_{t_1}/P^n_{t_0}$, which combines newly available goods, for which $p_{i_{t_0}}/p_{i_{t_1}} = 0$, and goods that were available in both periods but consumers did not have tastes for at t_0, for which $p_{i_{t_0}}/p_{i_{t_1}}$ is finite. In many applications, it is reasonable to think that $P^n_{t_1}/P^n_{t_0} \in [0, P^c_{t_1}/P^c_{t_0}]$. The lower bound assumes that newly consumed goods were priced at infinity in t_0. The upper bound assumes that changes in prices of newly consumed goods are not higher than changes in prices of continuing goods.

iii. The counterfactual share of exiting goods at t_0 prices but t_1 preferences $b^{x_{t_1}}_{t_0}$. It is reasonable to think that $b^{x_{t_1}}_{t_0} \in [0, b^{x_{t_0}}_{t_0}]$. The lower bound takes the position that exiting goods are no longer valuable to the consumer with t_1 tastes, and the upperbound takes the position that exiting are not relatively more valuable to the consumer with

[41] We can ignore goods that the consumer does not value at t_1 ($x_{it_1} = 0$) because those goods do not matter for welfare at t_1 preferences.

t_1 tastes than the consumer with t_0 tastes (i.e. demand curves for those goods that exited did not shift out).

By making different assumptions about the bounds, we can derive three noteworthy special cases of expression (25):

No extensive margin. Assume there are no new or exiting goods ($b_{t_1}^{\mathcal{C}} = 1$), so the price deflator is the same as in expression (9)

$$\frac{P_{t_1}}{P_{t_0}} = \frac{P_{t_1}^{\mathcal{C}}}{P_{t_0}^{\mathcal{C}}} = \left(\sum_{i \in \mathcal{C}} b_{it_1} \left(\frac{p_{it_1}}{p_{it_0}}\right)^{\theta_0 - 1}\right)^{\frac{1}{\theta_0 - 1}}. \tag{26}$$

This is the assumption we make in Section 5.2, since we only compute the price index for continuing goods.

Feenstra (1994) with taste shocks. Suppose that newly consumed goods at t_1 were unavailable at t_0 ($P_{t_1}^n / P_{t_0}^n = 0$). Furthermore, suppose that, under t_1 tastes but t_0 prices, the share of expenditures on exiting goods equals the observed share of exiting goods at t_0: ($b_{t_0}^{x_{t_1}} = b_{t_0}^{x_{t_0}} = 1 - b_{t_0}^c$). In this case, (25) reduces to

$$\frac{P_{t_1}}{P_{t_0}} = \left(\frac{b_{t_1}^c}{b_{t_0}^c}\right)^{\frac{1}{\theta_0 - 1}} \frac{P_{t_1}^c}{P_{t_0}^c}.$$

The term in front of the continuing price index coincides with the well-known new-goods adjustment with CES preferences due to Feenstra (1994). Note that P_{t_1} / P_{t_0} is still not the same as Feenstra (1994) because $P_{t_1}^c / P_{t_0}^c$ is calculated for fixed x_{t_1} tastes and hence is not the Sato-Vartia price index. In other words, $P_{t_1}^c / P_{t_0}^c$ is the price index in (26).

Entry/exit only due to taste shocks. Suppose that all goods were available in t_0 and t_1, but some goods are consumed in t_1 and not t_0 because of changes in tastes. In this case, $b_{t_0}^{x_{t_1}} = 0$, so

$$\frac{P_{t_1}}{P_{t_0}} = \left(b_{t_1}^c \left(\frac{P_{t_1}^c}{P_{t_0}^c}\right)^{\theta_0 - 1} + \left(1 - b_{t_1}^c\right) \left(\frac{P_{t_1}^n}{P_{t_0}^n}\right)^{\theta_0 - 1}\right)^{\frac{1}{\theta_0 - 1}}. \tag{27}$$

In this case, the price index is a weighted CES aggregator of the price index for continuing goods and the price index for newly consumed goods. The price index increases less than the price index for continuing goods if inflation for newly consumed goods is less than inflation for continuing goods.[42]

Beyond CES.

Our results in Section 4 can be generalized beyond CES functional forms relatively easily. In Appendix K, we discuss how Proposition 7 must be adjusted to allow for non-CES production and utility functions.

Heterogeneous agents.

Our microeconomic welfare results can be applied to individual households in economies without representative agents, but for brevity, we have abstracted from preference heterogeneity when defining and characterizing macroeconomic welfare. Baqaee and Burstein (2021) generalizes the results in this paper to environments where there is no representative agent using the Kaldor-Hicks compensation principle to measure societal welfare. Specifically, to measure the change in welfare from t_0 to t_1, we ask: *"what is the minimum amount endowments in t_0 must change so that it is possible to make every consumer indifferent between t_0 and t_1?"* We show that all the results in this paper generalize to economies without representative agents if we define social welfare in this way, and this definition collapses to the definition we use in this paper when there is a representative agent.

Inefficient Economies.

Our macro welfare results for neoclassical economies in Section 3 make use of the first welfare theorem. In Appendix L we show how to extend these results to economies with inefficiencies. We consider economies with distorting wedges between prices and marginal costs indexed by t. Our welfare measure EV^M is the proportional change in initial factor endowments so that the representative consumer with preferences $\succeq_{x_{t_1}}$ is indifferent between the economy with initial productivities and wedges and the economy with final productivities, endowments, and wedges.

Baqaee and Farhi (2019a) show that, in economies with stable and homothetic preferences, changes in welfare can be expressed in terms of the primitive shocks to technologies and wedges, Domar weights, and cost-based Domar weights. Appendix L shows that this

[42] For example, the model in Arkolakis (2016) predicts that changes in tastes induced by advertising will be correlated with changes in physical productivity, whereby more productive firms will expend more resources on advertising. In this case price of newly consumed goods increase on average less than continuing goods.

result generalizes to unstable and non-homothetic preferences if we use compensated Domar weights as in Definition 6, instead of observed ones.

7 Conclusion

In this paper, we provide a toolkit for studying how welfare changes in response to changes in budget sets or production possibility sets allowing for preference instability and non-homotheticity. In contrast to the case of stable and homothetic preferences, there is no model-free statistic like chain-weighted real consumption that simultaneously provides answers to different welfare questions (i.e. equivalent or compensating variation, initial or final preferences, partial or general equilibrium). We characterize the gap between real consumption, as measured in the data, and welfare-relevant objects. This gap can be large over long horizons relevant for long-run growth as well as for short-horizons, if demand-driven changes in expenditure shares covary with price or technology shocks.

Although our motivation and applications have focused on shocks across time, an interesting avenue to explore is welfare comparisons across locations (see e.g. Deaton, 2003, and Argente et al., 2021). Whereas in a temporal context, the preferences of today are more relevant than the preferences of yesterday, in a spatial context, both locations' preferences are equally interesting. The distinction between macroeconomic and microeconomic welfare is also important in a spatial context. Comparing budget constraints in one location to another may be misleading as a way to compare the technologies of two economies. This is because, even if PPFs in both locations are exactly the same, the relative price of goods households value more in one location will be lower in the other location.[43]

References

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

Acemoglu, D. and V. Guerrieri (2008). Capital deepening and nonbalanced economic growth. *Journal of political Economy 116*(3), 467–498.

Aghion, P., A. Bergeaud, T. Boppart, P. J. Klenow, and H. Li (2019). Missing growth from creative destruction. *American Economic Review 109*(8), 2795–2822.

Alder, S., T. Boppart, and A. Müller (2019). A theory of structural change that can fit the data.

[43] This implies that consumers from one location prefer the budget set in the other location, which resembles the Gerschenkron (1951) effect that the relative GDP of a country is higher when evaluated at another country's prices (the grass is greener in the other side). However, while the Gerschenkron effect is a spatial version of the discrepancy between Laspeyres and Paasche indices, our result is driven by the endogeneity of relative prices to demand forces and not by substitution bias.

Argente, D. O., C.-T. Hsieh, and M. Lee (2021). Measuring the cost of living in mexico and the us. *American Economic Journal: Macroeconomics*. Forthcoming.

Arkolakis, C. (2016). A unified theory of firm selection and growth. *The Quarterly Journal of Economics 131*(1), 89–155.

Atalay, E. (2017). How important are sectoral shocks? *American Economic Journal: Macroeconomics (Forthcoming)*.

Atkin, D., B. Faber, T. Fally, and M. Gonzalez-Navarro (2020, March). Measuring Welfare and Inequality with Incomplete Price Information. NBER Working Papers 26890, National Bureau of Economic Research, Inc.

Auer, R., A. Burstein, S. Lein, and J. Vogel (2021). Unequal expenditure switching: Evidence from switzerland.

Balk, B. M. (1989). Changing consumer preferences and the cost-of-living index: theory and nonparametric expressions. *Journal of Economics 50*(2), 157–169.

Baqaee, D. and A. Burstein (2021). Aggregate welfare and output with heterogeneous agents.

Baqaee, D. and E. Farhi (2020). Supply and demand in disaggregated keynesian economies with an application to the covid-19 crisis. Technical report, National Bureau of Economic Research.

Baqaee, D. R. and E. Farhi (2019a). Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics 135*(1), 105–163.

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

Barro, R. J. (2021, 01). Double Counting of Investment. *The Economic Journal 131*(638), 2333–2356.

Barrot, J.-N., B. Grassi, and J. Sauvagnat (2020). Sectoral effects of social distancing. *Available at SSRN*.

Basu, S., L. Pascali, F. Schiantarelli, and L. Serven (2012). Productivity and the welfare of nations. Technical report, National Bureau of Economic Research.

Baumol, W. J. (1967). Macroeconomics of unbalanced growth: the anatomy of urban crisis. *The American economic review*, 415–426.

Boehm, C., A. Flaaen, and N. Pandalai-Nayar (2015). Input linkages and the transmission of shocks: Firm-level evidence from the 2011 tōhoku earthquake.

Bonadio, B., Z. Huo, A. A. Levchenko, and N. Pandalai-Nayar (2020). Global supply chains in the pandemic. Technical report.

Boppart, T. (2014). Structural change and the kaldor facts in a growth model with relative price effects and non-gorman preferences. *Econometrica 82*(6), 2167–2196.

Buera, F. J. and J. P. Kaboski (2009, 04-05). Can Traditional Theories of Structural Change Fit The Data? *Journal of the European Economic Association 7*(2-3), 469–477.

Buera, F. J., J. P. Kaboski, and R. Rogerson (2015). Skill biased structural change. Technical report, National Bureau of Economic Research.

Cakmakli, C., S. Demiralp, S. Kalemli-Ozcan, S. Yesiltas, and M. A. Yildirim (2020, May). Covid-19 and emerging markets: An epidemiological multi-sector model for a small open economy with an application to turkey. Working Paper 27191, National Bureau of Economic Research.

Carvalho, V. and X. Gabaix (2013). The Great Diversification and its undoing. *The American Economic Review 103*(5), 1697–1727.

Carvalho, V. M. and A. Tahbaz-Salehi (2018). Production networks: A primer.

Cavallo, A. (2020). Inflation with covid consumption baskets. Technical report, National Bureau of Economic Research.

Caves, D. W., L. R. Christensen, and W. E. Diewert (1982). The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica: Journal of the Econometric Society*, 1393–1414.

Comin, D. A., D. Lashkari, and M. Mestieri (2015). Structural change with long-run income and price effects. Technical report, National Bureau of Economic Research.

Cravino, J., A. A. Levchenko, and M. Rojas (2019, September). Population Aging and Structural Transformation. NBER Working Papers 26327, National Bureau of Economic Research, Inc.

Deaton, A. (2003). Adjusted indian poverty estimates for 1999-2000. *Economic and political Weekly*, 322–326.

Deaton, A. and J. Muellbauer (1980). *Economics and consumer behavior*. Cambridge university press.

Diewert, W. E. (1976). Exact and superlative index numbers. *Journal of econometrics 4*(2), 115–145.

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

Fally, T. (2020). Integrability, generalized separability, and a new class of demand systems.

Feenstra, R. C. (1994). New product varieties and the measurement of international prices. *The American Economic Review*, 157–177.

Feenstra, R. C. and M. B. Reinsdorf (2000). An exact price index for the almost ideal demand system. *Economics Letters 66*(2), 159–162.

Feenstra, R. C. and M. B. Reinsdorf (2007, October). Should Exact Index Numbers Have Standard Errors? Theory and Application to Asian Growth. In *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, NBER Chapters, pp. 483–513. National Bureau of Economic Research, Inc.

Fisher, F. M. and K. Shell (1968). *Taste and quality change in the pure theory of the true cost of living index*. Value, Capital and Growth. Edinburgh University Press.

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

Gerschenkron, A. (1951). *A dollar index of Soviet machinery output, 1927-28 to 1937*, Volume 197. Rand Corporation.

Hamilton, B. W. (2001, June). Using engel's law to estimate cpi bias. *American Economic Review 91*(3), 619–630.

Hanoch, G. (1975). Production and demand models with direct or indirect implicit additivity. *Econometrica: Journal of the Econometric Society*, 395–419.

Hausman, J. (1981). Exact consumer's surplus and deadweight loss. *American Economic Review 71*(4), 662–76.

Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition. In *The economics of new goods*, pp. 207–248. University of Chicago Press.

Herrendorf, B., R. Rogerson, and A. Valentinyi (2013). Two perspectives on preferences and structural transformation. *American Economic Review 103*(7), 2752–89.

Hicks, J. R. (1940). The valuation of the social income. *Economica 7*(26), 105–124.

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.

IMF (2004). *Consumer Price Index Manual: Theory and Practice*. USA: International Monetary Fund.

Ivancic, L., W. E. Diewert, and K. J. Fox (2011). Scanner data, time aggregation and the construction of price indexes. *Journal of Econometrics 161*(1), 24–35.

Jaravel, X. (2019). The unequal gains from product innovations: Evidence from the us retail sector. *The Quarterly Journal of Economics 134*(2), 715–783.

Jones, C. I. and P. J. Klenow (2016). Beyond gdp? welfare across countries and time. *American Economic Review 106*(9), 2426–57.

Jorgenson, D. W., M. S. Ho, K. J. Stiroh, et al. (2005). productivity, volume 3: information technology and the american growth resurgence. *MIT Press Books 3*.

Kongsamut, P., S. Rebelo, and D. Xie (2001). Beyond Balanced Growth. *Review of Economic Studies 68*(4), 869–882.

Konüs, A. A. (1939). The problem of the true index of the cost of living. *Econometrica: Journal of the Econometric*

Society, 10–29.

Martin, R. S. (2020). Revisiting taste change in cost-of-living measurement. Technical report, US Department of Labor, US Bureau of Labor Statistics, Office of Prices and Living Conditions.

Matsuyama, K. (2019). Engel's law in the global economy: Demand-induced patterns of structural change, innovation, and trade. *Econometrica 87*(2), 497–528.

McKenzie, G. W. and I. F. Pearce (1982). Welfare measurement: a synthesis. *The American Economic Review 72*(4), 669–682.

Muellbauer, J. (1975). The cost of living and taste and quality change. *Journal of Economic Theory 10*(3), 269–283.

Ngai, L. R. and C. A. Pissarides (2007). Structural change in a multisector model of growth. *American economic review 97*(1), 429–443.

Nordhaus, W. D. et al. (2008). Baumol's diseases: A macroeconomic perspective. *The BE Journal of Macroeconomics 8*(1), 1–39.

Oulton, N. (2008). Chain indices of the cost-of-living and the path-dependence problem: An empirical solution. *Journal of Econometrics 144*(1), 306–324.

Redding, S. J. and D. E. Weinstein (2020). Measuring aggregate price indices with taste shocks: Theory and evidence for ces preferences. *The Quarterly Journal of Economics 135*(1), 503–560.

Samuelson, P. A. and S. Swamy (1974). Invariant economic index numbers and canonical duality: survey and synthesis. *The American Economic Review 64*(4), 566–593.

Syverson, C. (2017). Challenges to mismeasurement explanations for the us productivity slowdown. *Journal of Economic Perspectives 31*(2), 165–86.

Online Appendix

Appendix A Non-Parametric Comparison of EV^m and Real Consumption

In this appendix we provide and discuss an alternative version of Proposition 2 based on Hicksian demands without imposing non-homothetic CES preferences. The proof is in Appendix B.

Proposition 11 (Approximate Micro using Hicksian Demand)**.** *To a second-order approximation, the change in real consumption is*[44]

$$\Delta \log Y \approx \Delta \log I - b' \Delta \log p - \sum_{i \in N} \left[\frac{1}{2} \Delta \log p' \frac{\partial b_i}{\partial \log p} + \frac{1}{2} \Delta \log x' \frac{\partial b_i}{\partial \log x} + \frac{1}{2} \Delta \log v \frac{\partial b_i}{\partial \log u} \right] \Delta \log p,$$

and the change in welfare is

$$EV^m \approx \Delta \log I - b' \Delta \log p - \sum_{i \in N} \left[\frac{1}{2} \Delta \log p' \frac{\partial b_i}{\partial \log p} + \Delta \log x' \frac{\partial b_i}{\partial \log x} + \Delta \log v \frac{\partial b_i}{\partial \log u} \right] \Delta \log p,$$

where the derivatives are evaluated at t_0.

Comparing the expression for real consumption $\Delta \log Y$ and welfare EV^m shows that to a first order, they are the same. Discrepancies between the two arise starting at the second-order and involve how expenditure-switching is treated. Real consumption accounts for changes in budget shares in the same way regardless of their cause. The first term in the square brackets reflects changes in budget shares due to changes in relative prices (substitution effects) and the next two terms correspond to changes in budget share due to non-price factors (taste shocks and income effects).

The second line shows that welfare treats changes in budget shares due to substitution effects differently to changes in budget shares due to taste shocks or income effects. To understand the gap between welfare and real consumption changes, consider first the case of homothetic but unstable preferences. Whereas changes in real consumption only take into consideration changes in budget shares in response to taste shocks as the shock unfolds over time, changes in welfare account for these changes from the start. Therefore, changes in budget shares due to non-price factors are multiplied by $1/2$ in real consumption, but they are multiplied by 1 in welfare. In other words, real consumption does not

[44] The terms $\Delta \log x$ and $\Delta \log v$ need only be first-order approximations since they are multiplied by $\Delta \log p$ (and we only need to keep terms that are of order Δt^2). However, for the first term $-b\Delta \log p$, the primitive shock in prices must be approximated up to the second order, that is, $\Delta \log p \approx (\partial \log p / \partial t)\Delta t + 1/2(\partial \log p^2 / dt^2)\Delta t^2$.

sufficiently account for substitution caused by preference instability. For example, the additional reduction in welfare (at new preferences) from a price increase in a good i with increasing demand ($d\log x \frac{\partial b_i}{\partial \log x} d\log p_i > 0$) is not fully reflected in real consumption, implying $EV^m < \Delta \log Y$.

Similar reasoning applies in the case of stable but non-homothetic preferences, since changes in budget shares due to non-homotheticities should be incorporated in welfare immediately but are reflected in real consumption only gradually. For example, a reduction in the price of a good for which income effects are relatively weak ($d\log v \frac{\partial b_i}{\partial \log v} d\log p_i > 0$) implies a smaller increase in welfare than in real consumption ($EV^m < \Delta \log Y$).

Appendix B Proofs

Proof of Lemma 1. By definition,

$$\begin{aligned} EV^m &= \log \frac{e\left(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1}\right)}{e\left(p_{t_0}, v(p_{t_0}, I_{t_0}; x_{t_1}); x_{t_1}\right)} \\ &= \log \frac{e\left(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1}\right)}{e\left(p_{t_0}, v(p_{t_0}, I_{t_0}; x_{t_1}); x_{t_1}\right)} \frac{e\left(p_{t_1}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1}\right)}{e\left(p_{t_1}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1}\right)} \\ &= \log \frac{e\left(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1}\right)}{I_{t_0}} \frac{I_{t_1}}{e\left(p_{t_1}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1}\right)}. \end{aligned}$$

To finish, rewrite

$$\log \frac{e\left(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1}\right)}{e\left(p_{t_1}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1}\right)} = -\int_{t_0}^{t_1} \sum_{i\in N} \frac{\partial \log e(p_t, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})}{\partial \log p_{it}} \frac{d\log p_{it}}{dt} dt,$$

and use the Shephard's lemma to express the price elasticity of the expenditure function in terms of budget shares. If the path of prices between t_0 and t_1 is not differentiable, then construct a new modified path of prices that is differentiable, and apply the integral to this modified path. Since the integral is path independent, it only depends on p_{t_0} and p_{t_1}. Therefore any smooth path that connects p_{t_0} and p_{t_1} gives the same integral. □

Proof of Proposition 1. If the path of prices is continuously differentiable, we can combine Lemma 1 with the definition of real consumption. □

Proof of Proposition 11. For real consumption, differentiate real consumption

$$\Delta \log Y = \int_{t_0}^{t_1} \left(\frac{d\log I_t}{dt} dt - \sum_{i\in N} b_i(p_t, u_t; x_t) \frac{d\log p_{it}}{dt} \right) dt$$

twice with respect to t_1 and evaluate the derivative at $t_1 = t_0$. This yields the desired expression.

For EV^m, using Lemma 1, we can write

$$EV^m = \Delta \log I - \int_{t_0}^{t_1} \sum_{i \in N} \frac{\partial \log e(p, v(p_{t_1}, x_{t_1}), x_{t_1})}{\partial \log p_i} \frac{d \log p_{it}}{dt} dt$$

Differentiate EV^m twice with respect to t_1 and evaluate the derivative at $t_1 = t_0$

$$\begin{aligned}
\frac{dEV^m}{dt_1} &= \frac{d \log I}{dt} - \sum_{i \in N} \frac{\partial \log e(p, v(p_{t_1}, x_{t_1}), x_{t_1})}{\partial \log p_i} \frac{d \log p_{it}}{dt} \\
&- \int_{t_0}^{t_1} \sum_{i \in N} d \log v \frac{\partial^2 \log e(p, v(p_{t_1}, x_{t_1}), x_{t_1})}{\partial \log u \partial \log p_i} d \log p_{it} \\
&- \int_{t_0}^{t_1} \sum_{i \in N} d \log x \frac{\partial^2 \log e(p, v(p_{t_1}, x_{t_1}), x_{t_1})}{\partial \log x \partial \log p_i} d \log p_{it} \\
\frac{d^2 EV^m}{dt_1^2} &= \frac{d^2 \log I}{dt^2} - \sum_{i \in N} b_i \frac{d^2 \log p_{it}}{dt^2} - \sum_{i \in N} \sum_{j \in N} \frac{\partial^2 \log e(p, v(p_{t_1}, x_{t_1}), x_{t_1})}{\partial \log p_i \partial \log p_j} d \log p_{it} d \log p_{jt} \\
&- 2 \sum_{i \in N} d \log v_t \frac{\partial^2 \log e(p, v(p_{t_1}, x_{t_1}), x_{t_1})}{\partial \log p_i \partial \log u} d \log p_{it} - 2 \sum_{i \in N} d \log x_t' \frac{\partial \log e(p, v(p_{t_1}, x_{t_1}), x_{t_1})}{\partial \log p_{it} \partial \log x} d \log p_{it} \\
&= \frac{d^2 \log I}{dt^2} - \sum_{i \in N} \sum_{j \in N} \frac{\partial b_i}{\partial \log p_j} d \log p_{it} d \log p_{jt} \\
&- \sum_{i \in N} b_i \frac{d^2 \log p_{it}}{dt^2} - 2 \sum_{i \in N} d \log v \frac{\partial b_i}{\partial \log u} d \log p_{it} - 2 \sum_{i \in N} d \log x_t' \frac{\partial b_i}{\partial \log x} d \log p_{it} \\
&= \frac{d^2 \log I}{dt^2} - \sum_{i \in N} \left[\sum_{j \in N} \frac{\partial b_i}{\partial \log p_j} d \log p_{jt} + d \log v_t \frac{\partial b_i}{\partial \log u} + d \log x_t' \frac{\partial b_i}{\partial \log x} \right] d \log p_{it} \\
&- \sum_{i \in N} b_i \frac{d^2 \log p}{dt^2} - \sum_{i \in N} d \log v_t \frac{\partial b_i}{\partial \log u} d \log p_{it} - \sum_{i \in N} d \log x_t' \frac{\partial b_i}{\partial \log x} d \log p_{it} \\
&= \frac{d^2 \log I}{dt^2} - \sum_{i \in N} d b_{it} d \log p_{it} - \sum_{i \in N} b_i \frac{d^2 \log p_{it}}{dt^2} \\
&- \sum_{i \in N} d \log v \frac{\partial b_i}{\partial \log u} d \log p_{it} - \sum_{i \in N} d \log x_t' \frac{\partial b_i}{\partial \log x} d \log p_{it}
\end{aligned}$$

The first three terms are equal to the second-order expansion of $\Delta \log Y$, and the remaining terms are the bias. □

Proof of Proposition 2. By Proposition 11, we have

$$\Delta \log Y \approx \Delta \log I - \sum_i b_i \Delta \log p_i - \frac{1}{2} \sum_i \Delta b_i \Delta \log p_i .$$

Substitute (7) in place of Δb to get the desired expression. For the gap between real consumption and EV^m, note that Proposition 1 implies that

$$EV^m - \Delta \log Y \approx -\frac{1}{2} \sum_i \left[\Delta b_i - \sum_j \frac{\partial b_i^H}{\partial \log p_j} \Delta \log p_j \right] \Delta \log p_i$$

where b^H is the Hicksian budget share (holding fixed utility and demand shifters). Using (7) in place of Δb above and the fact that $\frac{\partial b_i^H}{\partial \log p_i} = (1-\theta_0) b_i (1-b_i)$ for $i = j$ and $\frac{\partial b_i^H}{\partial \log p_j} = \theta_0 b_i b_j$ for $i \neq j$, yields the following

$$\Delta \log EV^m - \Delta \log Y \approx -\frac{1}{2} \sum_{i \in N} \left[(\varepsilon_i - 1) b_i \left(d \log I - \sum_{j \in N} b_j \Delta \log p_j \right) + b_i \Delta \log x_i \right] \Delta \log p_i ,$$

which can be rearranged to give the desired expression. □

Proof of Proposition 4. For this proof, we use notation introduced in Section 4. Setting nominal GDP to be the numeraire, we can write (simplifying notation by suppressing dependence on t in the integral)

$$\begin{aligned}
\Delta \log Y &= -\int_{t_0}^{t_1} b' d \log p \\
&= -\int_{t_0}^{t_1} b' \left[-\Psi d \log A - \Psi^F d \log L + \Psi^F d \log \Lambda \right] \\
&= \int_{t_0}^{t_1} b' \Psi d \log A - \int_{t_0}^{t_1} b' \Psi^F [d \log \Lambda - d \log L] \\
&= \int_{t_0}^{t_1} \lambda' d \log A + \int_{t_0}^{t_1} \Lambda' d \log L - \int_{t_0}^{t_1} \Lambda d \log \Lambda \\
&= \int_{t_0}^{t_1} \lambda' d \log A + \int_{t_0}^{t_1} \Lambda' d \log L
\end{aligned}$$

where the second line uses Proposition 7, and we use the fact that $\lambda' = b' \Psi$, $\Lambda' = b' \Psi^F$, and $b' \Psi^F d \log \Lambda = \Lambda' d \log \Lambda = 0$ because the factor shares always sum to one: $\sum_{f \in F} \Lambda_f = 1$. □

Proof of Proposition 5. Recall that the macro equivalent variation at final preferences is defined by $EV^M = \phi$, where

$$V\left(A_{t_0}, e^{\phi} L_{t_0}; x_{t_1}\right) = V\left(A_{t_1}, L_{t_1}; x_{t_1}\right)$$

Denote by $p\left(A, L, x\right)$ goods prices under technologies A, factor quantities L, and preferences x. Without loss of generality, we fix income at I. We have $p_{t_1} \equiv p\left(A_{t_1}, L_{t_1}, x_{t_1}\right)$ and

$$v_{t_1} \equiv v\left(p_{t_1}, I; x_{t_1}\right) = V\left(A_{t_1}, L_{t_1}; x_{t_1}\right).$$

Define a hypothetical economy with fictional households that have stable homothetic preferences defined by the expenditure function $e^{ev}\left(p, u\right) = e\left(p, v_{t_1}; x_{t_1}\right) u$. Budget shares of this fictional consumer are $b_i^{ev}\left(p\right) \equiv \frac{\partial \log e^{ev}(p,u)}{\partial \log p_i} = \frac{\partial \log e\left(p, v_{t_1}; x_{t_1}\right)}{\partial \log p_i}$. Given any technology vector, in this hypothetical economy we denote the Leontief inverse matrix by Ψ^{ev} and sales shares by λ^{ev}. Given technologies A_t and factor quantities L_t, we denote prices in this hypothetical economy by p_t^{ev}. Changes in prices in this hypothetical economy satisfy

$$d \log p^{ev} = -\Psi^{ev} d \log A + \Psi^{evF} d \log \Lambda^{ev}, \tag{28}$$

where Ψ^{ev} is the fictitious Leontief inverse. Note that $p\left(A_{t_1}, L_{t_1}, x_{t_1}\right) = p^{ev}\left(A_{t_1}, L_{t_1}\right)$ and $p\left(A_{t_0}, e^{\phi} L_{t_0}, x_{t_1}\right) = p^{ev}\left(A_{t_0}, e^{\phi} L_{t_0}\right)$, where we used the fact that $V\left(A_{t_0}, e^{\phi} L_{t_0}; x_{t_1}\right) = v_{t_1}$. We will use the property that, with constant returns to scale, homothetic preferences, and constant income I,

$$p^{ev}\left(A, aL\right) = \frac{1}{a} p^{ev}\left(A, L\right)$$

for every $a > 0$. Using the previous results,

$$\begin{aligned} V\left(A_{t_0}, e^{\phi} L_{t_0}; x_{t_1}\right) &= v\left(p\left(A_{t_0}, e^{\phi} L_{t_0}, x_{t_1}\right), I; x_{t_1}\right) \\ &= v\left(p^{ev}\left(A_{t_0}, e^{\phi} L_{t_0}\right), I; x_{t_1}\right) \\ &= v\left(e^{-\phi} p^{ev}\left(A_{t_0}, L_{t_0}\right), I; x_{t_1}\right) \\ &= v\left(p^{ev}\left(A_{t_0}, L_{t_0}\right), e^{\phi} I; x_{t_1}\right), \end{aligned}$$

where the last equality used the fact that the value function is homogeneous of degree 0 in prices and income. We thus have

$$v\left(p^{ev}\left(A_{t_0}, L_{t_0}\right), e^{\phi} I; x_{t_1}\right) = v\left(p^{ev}\left(A_{t_1}, L_{t_1}\right), I; x_{t_1}\right),$$

which can be re-expressed using the expenditure function as

$$EV^M = \log \frac{e\left(p^{ev}\left(A_{t_1}, L_{t_1}\right), v_{t_1}; x_{t_1}\right)}{e\left(p^{ev}\left(A_{t_0}, L_{t_0}\right), v_{t_1}; x_{t_1}\right)}.$$

This observation is a key step in the proof. Macro welfare changes can be re-expressed as micro welfare changes given changes in equilibrium prices in a fictional economy with preferences represented by $e^{ev}(p, u)$. As in the proof of Lemma 1, rewrite EV^M as (suppressing dependence on t in the integral)

$$EV^M = -\int_{t_0}^{t_1} \sum_{i \in N+F} \frac{\partial \log e(p, v_{t_1})}{\partial \log p_i} d\log p_i^{ev} = -\int_{t_0}^{t_1} \sum_{i \in N+F} b_i^{ev} d\log p_i^{ev}.$$

Following the same steps as in the proof of Proposition 4 (for the hypothetical economy), we obtain

$$EV^M = \int_{t_0}^{t_1} \sum_{i \in N} \lambda_i^{ev} d\log A_i + \int_{t_0}^{t_1} \sum_{f \in F} \lambda_i^{ev} d\log L_f.$$

□

In general, macro and micro welfare changes are not the same when preferences are unstable and nonhomothetic. However, when the PPF is linear, the following proposition shows that they coincide.

Proof of Proposition 3. By the proof of Proposition 5, $EV^m = EV^M$ if and only if $p^{ev}(A_t, L_t) = p(A_t, L_t, x_t)$ for all t. This condition is immediate if preferences are homothetic and stable. Consider now the case in which preferences are non-homothetic and/or unstable but factor income shares, Λ, are constant. Then by Proposition 7, changes in prices in response to changes in A, L, and x are given by the following differential equation:

$$d\log p = -\Psi d\log A - \Psi^F d\log L.$$

Furthermore, note that changes in Ψ are determined by changes in Ω since $\Psi = (I - \Omega)^{-1}$. Since every $i \in N$ has constant returns to scale, changes in Ω_{ij} depend only on changes in relative prices for every $i \in N$. This means that changes in Ω only depend on changes in relative prices, therefore changes in Ψ depend only on changes in relative prices. Since x and utility v do not appear in any of these expressions, this means that prices and incomes $p(A, L, x)$ and $I(A, L, x)$, relative to the numeraire, do not depend on x and v. Thus, $p^{ev}(A_t, L_t) = p(A_t, L_t, x_t)$.

□

Proof of Proposition 6. Differentiate real GDP (abstracting from changes in factor endowments),

$$\Delta \log Y = \int_{t_0}^{t_1} \sum_{i \in N} \lambda_i(A(t), u(t); x(t)) \frac{d \log A_i}{dt} dt,$$

twice with respect to t_1 and evaluate the derivative at $t_1 = t_0$. This yields the desired expression. Following similar steps as in the proof of Proposition 2,

$$EV^M \approx \Delta \log Y + \frac{1}{2} \sum_{i \in N} \left[\Delta \lambda_i - \sum_{j \in N} \frac{\partial \lambda_i^{ev}}{\partial \log A_j} \Delta \log A_j \right] \Delta \log A_i.$$

The term in square brackets is the change in sales shares due to changes in utility and demand shifters. This expression can be written as

$$EV^M \approx \Delta \log Y + \frac{1}{2} \sum_{i \in N} \left[\Delta \log x' \frac{\partial \lambda_i}{\partial \log x} + \Delta \log V \frac{\partial \lambda_i}{\partial \log u} \right] \Delta \log A_i. \tag{29}$$

□

Proof of Proposition 7. We normalize nominal GDP to be the numeraire. Then Shephard's lemma implies that, for each $i \in N$

$$d \log p_i = -d \log A_i + \sum_{j \in N+F} \Omega_{ij} d \log p_j,$$

where $d \log p_j$ is the change in the price of $j \in N + F$. For $i \in F$

$$d \log p_i = -d \log A_i + d \log \Lambda_i.$$

Combining these yields the desired expression for changes in prices

$$d \log p = -\Psi d \log A + \Psi^F d \log \Lambda.$$

To get changes in sales shares, note that

$$\begin{aligned} \lambda &= b'\Psi \\ d\lambda &= d(b'\Psi) \\ &= b'\Psi d\Omega \Psi + db'\Psi \end{aligned}$$

$$
\begin{aligned}
\Omega_{ij} d\log\Omega_{ij} &= (1-\theta_i)\Omega_{ij}(d\log p_j - \sum_k \Omega_{ik} d\log p_k) \\
d\Omega_{ij} &= (1-\theta_i) Cov_{\Omega_{(i,:)}}(d\log p, Id_{(:,j)}) \\
\sum_j d\Omega_{ij}\Psi_{jk} &= (1-\theta_i) Cov_{\Omega_{(i,:)}}(d\log p, Id_{(:,j)})\Psi_{jk} \\
&= (1-\theta_i) \sum_j Cov_{\Omega_{(i,:)}}(d\log p, \Psi_{jk} Id_{(:,j)}) \\
[d\Omega\Psi]_{ik} &= (1-\theta_i) Cov_{\Omega_{(i,:)}}(d\log p, \Psi_{(:,k)})
\end{aligned}
$$

Meanwhile

$$
\begin{aligned}
d\log b_i &= (1-\theta_0)\left(d\log p_i - \sum_i b_i d\log p_i\right) + (\varepsilon_i - 1) d\log Y + d\log x_i \\
&= (1-\theta_0) Cov_{\Omega_{(0,:)}}\left(d\log p, Id_{(:,i)}\right) + Cov_{\Omega_{(0,:)}}\left(\varepsilon, Id_{(:,i)}\right) d\log Y + Cov_{\Omega_{(0,:)}}\left(d\log x, Id_{(:,i)}\right) \\
\sum_i db_i \Psi_{ik} &= Cov_{\Omega_{(0,:)}}\left((1-\theta_0) d\log p + \varepsilon d\log Y + d\log x, \Psi_{(:,k)}\right)
\end{aligned}
$$

Hence, $d\lambda' = \lambda' d\Omega\Psi + db'\Psi$ can be written as

$$
d\lambda_k = \sum_i \lambda_i (1-\theta_i) Cov_{\Omega_{(i,:)}}(d\log p, \Psi_{(:,k)}) + Cov_{\Omega_{(0,:)}}\left(\varepsilon, \Psi_{(:,k)}\right) d\log Y + Cov_{\Omega_{(0,:)}}\left(d\log x, \Psi_{(:,k)}\right).
$$

□

Proof of Proposition 8. Normalize nominal GDP to one. Applying Proposition 7 to a one-factor model yields

$$
d\log p = -\Psi d\log A,
$$

so that relative prices do not respond to changes in demand or income.

To solve for $\Delta \log Y$, use Proposition 6 in combination with the expression for $d\log p$ and $d\lambda$ in Proposition 7 in the case of one factor. To solve for EV^M, by Proposition 3, $EV^M = EV^m$. Solve for $EV^m - \Delta\log Y$ by plugging the expression for $d\log p$ into Proposition 11 and noting that $b' = \Omega_{(0,:)}$. □

Proof of Proposition 9. Consider intertemporal preferences

$$
V(\boldsymbol{A}, \boldsymbol{L}, K_0) = \sum_{s=t}^{\infty} \beta^{s-t} u(C_s).
$$

Comparing economies t and t', macro EV solves the following equation:

$$V(\boldsymbol{A}, \phi \boldsymbol{L}, \phi K_0) = \sum_{s=t}^{\infty} \beta^{s-t} u(C_s\left(\boldsymbol{A}, \phi \boldsymbol{L}, \phi K_0\right)) = \sum_{s=t'}^{\infty} \beta^{s-t'} u(C_s\left(\boldsymbol{A}', \boldsymbol{L}', K_0'\right)) = V(\boldsymbol{A}', \boldsymbol{L}', K_0').$$

Since the economy t' is in steady-state, we are looking for

$$\sum_{s=t}^{\infty} \beta^{s-t} u\left(C_s\left(\boldsymbol{A}, \phi \boldsymbol{L}, \phi K_0\right)\right) = \frac{1}{1-\beta} u\left(C\left(\boldsymbol{A}', \boldsymbol{L}', K_0'\right)\right).$$

Furthermore, since $(\boldsymbol{A}, \phi \boldsymbol{L}, \phi K_0)$ is also a steady-state (by Lemma 2 below), we are searching for

$$u\left(C\left(\boldsymbol{A}, \phi \boldsymbol{L}, \phi K_0\right)\right) = u\left(C\left(\boldsymbol{A}', \boldsymbol{L}', K_0'\right)\right)$$

or

$$C\left(\boldsymbol{A}, \phi \boldsymbol{L}, \phi K_0\right) = C\left(\boldsymbol{A}', \boldsymbol{L}', K_0'\right).$$

Let $v(\boldsymbol{p}, I)$ be the static indirect utility function. Then we know that we are searching for

$$v(p(\boldsymbol{A}, \phi \boldsymbol{L}, \phi K_0), m) = v(p(\boldsymbol{A}, \boldsymbol{L}, K_0), \phi m) = v(p(\boldsymbol{A}', \boldsymbol{L}', K_0'), m'),$$

where the first equality uses the fact within period relative goods prices do not depend on within period preferences (since the static PPF is linear). Hence,

$$\begin{aligned}
\phi &= \frac{e(p(\boldsymbol{A}, \boldsymbol{L}, K_0), v_{t_1})}{e(p(\boldsymbol{A}, \boldsymbol{L}, K_0), v_{t_0})} = \frac{e(p(\boldsymbol{A}, \boldsymbol{L}, K_0), v_{t_1})}{e(p(\boldsymbol{A}, \boldsymbol{L}, K_0), v_{t_0})} \frac{e(p(\boldsymbol{A}', \boldsymbol{L}', K_0'), v_{t_1})}{e(p(\boldsymbol{A}', \boldsymbol{L}', K_0'), v_{t_1})} \\
&= \frac{e(p(\boldsymbol{A}', \boldsymbol{L}', K_0'), v_{t_1})}{e(p(\boldsymbol{A}, \boldsymbol{L}, K_0), v_{t_0})} \frac{e(p(\boldsymbol{A}, \boldsymbol{L}, K_0), v_{t_1})}{e(p(\boldsymbol{A}', \boldsymbol{L}', K_0'), v_{t_1})} \\
&= \exp EV^m.
\end{aligned}$$

Hence, we can use micro EV^m to calculate the change in macro welfare. □

Lemma 2. *The steady-state choice of capital (and investment) is the same for any homothetic and stable within-period preferences.*

Proof. Suppose intertemporal welfare is given by

$$U_t = \sum_{s=t}^{\infty} \beta^{s-t} u(C_s),$$

where C_s is some homothetic aggregator of within-period consumption goods. Since all goods are produced with constant-returns to scale and every good uses the same homothetic bundle of capital and labor, we can write the consumption aggregator as depending on

$$C_s = G(L_{cs}, K_{cs})$$

for some function constant-returns-to-scale function G. Similarly, investment goods are created according to some constant returns to scale function

$$X_s = X(L_{Xs}, K_{Xs}),$$

and the capital accumulation equation is

$$K_{s+1} = (1-\delta)(K_s + X_s).$$

The Lagrangean is

$$\mathcal{L} = \sum_{s=t}^{\infty} \beta^{s-t} \left[u(C_s) + \mu_s(G(L_{cs}, K_{cs}) - C_s) + \kappa_s(K_{s+1} - (1-\delta)(K_s + X(L_{Xs}, K_{Xs}))) \right.$$
$$\left. + \rho_s(L_s - L_{cs} - L_{Xs}) + \psi_t(K_s - K_{cs} - K_{Xs}) \right]$$

The first order conditions are

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial C_s} &: u'(C_s) = \mu_s \\
\frac{\partial \mathcal{L}}{\partial K_{s+1}} &: \kappa_s - \beta\kappa_{s+1}(1-\delta) + \beta\psi_{s+1} = 0 \\
\frac{\partial \mathcal{L}}{\partial K_{Xs}} &: -\kappa_s(1-\delta)\frac{\partial X_s}{\partial K_{Xs}} = \psi_s = \mu_s \frac{\partial G}{\partial K_{cs}} \\
\frac{\partial \mathcal{L}}{\partial K_{cs}} &: \mu_s \frac{\partial G}{\partial K_{cs}} = \psi_s \\
\frac{\partial \mathcal{L}}{\partial L_{cs}} &: \mu_s \frac{\partial G}{\partial L_{cs}} = \rho_s \\
\frac{\partial \mathcal{L}}{\partial L_{Xs}} &: -\kappa_s(1-\delta)\frac{\partial X}{\partial L_{Xs}} = \rho_s.
\end{aligned}$$

Hence

$$-\kappa_s(1-\delta) = \mu_s \frac{\partial G / \partial K_{cs}}{\partial X_s / \partial K_{Xs}}$$

$$\begin{aligned}
\kappa_s &= \beta\kappa_{s+1}(1-\delta) - \beta\psi_{s+1} \\
u'(C_s) &= \beta(1-\delta)u'(C_{s+1})\frac{\partial G/\partial K_{cs+1}}{\partial G/\partial K_{cs}} \partial X_s/\partial K_{Xs} \left[(\partial X_s/\partial K_{Xs+1})^{-1} + 1 \right].
\end{aligned}$$

In steady state we have

$$1 = \beta(1-\delta)\left[1 + \partial X_s / \partial K_{Xs}\right].$$

Hence, the capital stock and investment in steady-state are pinned down by the following 5 equations in 5 unknowns (K_C, K_X, K, L_C, L_I) :

$$\begin{aligned}
1 &= \beta(1-\delta)\left[1 + \partial X / \partial K_X\right], \\
\frac{K_C}{L_C} &= \frac{K_X}{L_X}, \\
K &= K_C + K_X, \\
L &= L_C + L_X, \\
\delta K &= (1-\delta) X(L_X, K_X).
\end{aligned}$$

Since G does not appear in any of these equations, the steady-state investment and capital stock do not depend on the shape of the within-period utility function G. □

Proof of Proposition 10. By Lemma 1, welfare changes measured as the equivalent variation at t_1 preferences is given by[45]

$$EV^m = \log \frac{I_{t_1}}{I_{t_0}} - \log \frac{P_{t_1}}{P_{t_0}}, \tag{30}$$

where

$$\frac{P_{t_1}}{P_{t_0}} = \left(\frac{\sum_i x_{it_1} p_{it_0}^{1-\theta_0}}{\sum_j x_{jt_1} p_{jt_1}^{1-\theta_0}} \right)^{\frac{1}{\theta_0 - 1}} = \left(\frac{\sum_i x_{it_1} p_{it_1}^{1-\theta_0} \left(\frac{p_{it_0}}{p_{it_1}}\right)^{1-\theta_0}}{\sum_j x_{jt_1} p_{jt_1}^{1-\theta_0}} \right)^{\frac{1}{\theta_0 - 1}}. \tag{31}$$

We re-express (31) as

$$\begin{aligned}
\left(\frac{P_{t_1}}{P_{t_0}}\right)^{\theta_0 - 1} &= \frac{\sum_{i\in\mathcal{C}} x_{it_1} p_{it_1}^{1-\theta_0} \left(\frac{p_{it_0}}{p_{it_1}}\right)^{1-\theta_0} + \sum_{i\in\mathcal{N}} x_{it_1} p_{it_1}^{1-\theta_0} \left(\frac{p_{it_0}}{p_{it_1}}\right)^{1-\theta_0} + \sum_{i\in\mathcal{X}} x_{it_1} p_{it_0}^{1-\theta_0}}{\sum_j x_{jt_1} p_{jt_1}^{1-\theta_0}} \\
&= b^c_{t_1} \sum_{i\in\mathcal{C}} b^c_{it_1} \left(\frac{p_{it_0}}{p_{it_1}}\right)^{1-\theta_0} + \left(1 - b^c_{t_1}\right) \sum_{i\in\mathcal{N}} b^n_{it_1} \left(\frac{p_{it_0}}{p_{it_1}}\right)^{1-\theta_0} + b^{x_{t_1}}_{t_0} \left(\frac{P_{t_1}}{P_{t_0}}\right)^{\theta_0 - 1}
\end{aligned} \tag{32}$$

where $b^c_{t_1} \equiv \sum_{i\in\mathcal{C}} b_{it_1}$ denotes the expenditure share on continuing goods at t_1,

$$b^c_{it_1} \equiv \frac{x_{it_1} p_{it_1}^{1-\theta_0}}{\sum_{i\in\mathcal{C}} x_{it_1} p_{it_1}^{1-\theta_0}}, \qquad b^n_{it_1} \equiv \frac{x_{it_1} p_{it_1}^{1-\theta_0}}{\sum_{i\in\mathcal{N}} x_{it_1} p_{it_1}^{1-\theta_0}},$$

[45] With homothetic preferences, equivalent and compensating variation are the same.

and

$$b_{t_0}^{x_{t_1}} \equiv \frac{\sum_{i\in\mathcal{X}} x_{it_1} p_{it_0}^{1-\theta_0}}{\sum_j x_{jt_1} p_{jt_0}^{1-\theta_0}}$$

is the (unobserved) share of exiting goods at t_0 prices under t_1 preferences. Defining a price index for continuing goods under t_1 preferences,

$$\frac{P_{t_1}^c}{P_{t_0}^c} = \left(\sum_{i\in\mathcal{C}} b_{it_1}^c \left(\frac{p_{it_1}}{p_{it_0}}\right)^{\theta_0-1}\right)^{\frac{1}{\theta_0-1}}$$

and a price index for new goods under t_1 preferences,

$$\frac{P_{t_1}^n}{P_{t_0}^n} = \left(\sum_{i\in\mathcal{N}} b_{it_1}^n \left(\frac{p_{it_1}}{p_{it_0}}\right)^{\theta_0-1}\right)^{\frac{1}{\theta_0-1}},$$

we can express the change in the welfare-relevant price index as

$$\frac{P_{t_1}}{P_{t_0}} = \left(\frac{b_{t_1}^c \left(\frac{P_{t_1}^c}{P_{t_0}^c}\right)^{\theta_0-1} + \left(1-b_{t_1}^c\right)\left(\frac{P_{t_1}^n}{P_{t_0}^n}\right)^{\theta_0-1}}{1-b_{t_0}^{x_{t_1}}}\right)^{\frac{1}{\theta_0-1}} \tag{33}$$

□

Appendix C Extension to Other Welfare Measures

Our baseline measure of welfare changes is equivalent variation under final preferences. Alternatively, we could measure changes in welfare using compensating (instead of equivalent) variation, or by using initial (rather than final) preferences. In this appendix, we show how to calculate the alternative welfare measures. Note that if preferences are homothetic, then the expenditure function can be written as $e(p,u;x) = e(p;x)\,u$, so for any x equivalent and compensating variation are equal. If preferences are stable, then the expenditure function can be written as $e(p,u;x) = e(p,u)$, so equivalent variation under initial and final preferences are equal (and the same is the case for compensating variation).

Micro welfare changes We consider four alternative measures of micro welfare changes.

The *compensating variation* with *initial preferences* is $CV^m(p_{t_0}, I_{t_0}, p_{t_1}, I_{t_1}; x_{t_0}) = \phi$, where

ϕ solves

$$v(p_{t_1}, e^{-\phi} I_{t_1}; x_{t_0}) = v(p_{t_0}, I_{t_0}; x_{t_0}). \tag{34}$$

The analog to (5) in Lemma 1 is (suppressing dependence on t in the integral)

$$CV^m = \Delta \log I - \int_{t_0}^{t_1} \sum_{i \in N} b_{it}^{cv} d \log p_{it}, \tag{35}$$

where $b_{it}^{cv} \equiv b_i(p_t, v(p_{t_0}, I_{t_0}; x_{t_0}); x_{t_0})$.

Whereas EV^m weights price changes by hypothetical budget shares evaluated at current prices for fixed *final preferences and final utility*, CV^m uses budget shares evaluated at current prices for fixed *initial preferences and initial utility*. An alternative way of calculating CV^m is to reverse the flow of time (the final period corresponds to the initial period), calculate the baseline EV measure under this alternative timeline, and then set $CV^m = -EV^m$.

We now briefly describe how to calculate b^{cv} to apply (35). For ex-ante counterfactuals, where b_{t_0} is known, we can construct $b^{cv}(p)$ between t_0 and t_1 by iterating on (8) (with b^{cv} in place of b^{ev}) starting at t_0 and going forward to t_1. For ex-post counterfactuals, b_{t_0} can be obtained from past data, so we can construct $b^{cv}(p)$ by iterating on (8) starting at t_0 and going forward to t_1. For non-homothetic CES,

$$CV^m = \Delta \log I - \log \left(\sum_i b_{it_0} \left(\frac{p_{it_1}}{p_{it_0}} \right)^{1-\theta_0} \right)^{\frac{1}{1-\theta_0}}. \tag{36}$$

To a second-order approximation around t_0 (as in Proposition 11)

$$\begin{aligned} \Delta \log CV^m &\approx \Delta \log I - b' \Delta \log p - \frac{1}{2} \sum_{i \in N} \left[\Delta \log p' \frac{\partial b_i}{\partial \log p} \right] \Delta \log p \\ &\approx \Delta \log Y + \frac{1}{2} \sum_{i \in N} \left[\Delta \log x' \frac{\partial b_i}{\partial \log x} + \Delta \log v \frac{\partial b_i}{\partial \log u} \right] \Delta \log p. \end{aligned} \tag{37}$$

Recall that changes in budget shares due to non-price factors are multiplied by $1/2$ in real consumption. However, they are multiplied by 0 in CV^m, since CV^m is based on budget shares at initial preferences and initial utility.

Combining Proposition 11 and (37), we see that up to a second order approximation,

$$0.5 \left(EV^m + CV^m \right) \approx \Delta \log Y. \tag{38}$$

That is, locally (but not globally) changes in real consumption equal a simple average of equivalent variation under final preferences and compensating variation under initial preferences.

Alternatively, we can measure the change in welfare using the *micro equivalent variation* with *initial preferences*, $EV^m(p_{t_0}, I_{t_0}, p_{t_1}, I_{t_1}; x_{t_0}) = \phi$ where ϕ solves

$$v(p_{t_1}, I_{t_1}; x_{t_0}) = v(p_{t_0}, e^{\phi} I_{t_0}; x_{t_0}). \tag{39}$$

Globally, changes in welfare are given by (5) where $b_{it}^{ev} \equiv b_i(p_t, v(p_{t_1}, I_{t_1}; x_{t_0}); x_{t_0})$. Finally, the change in welfare measured using the *micro compensating variation* with *final preferences* is $CV^m(p_{t_0}, I_{t_0}, p_{t_1}, I_{t_1}; x_{t_1}) = \phi$ where ϕ solves

$$v(p_{t_1}, e^{-\phi} I_{t_1}; x_{t_1}) = v(p_{t_0}, I_{t_0}; x_{t_1}). \tag{40}$$

Globally, changes in welfare are given by (35) where $b_{it}^{cv} \equiv b_i(p_t, v(p_{t_0}, I_{t_0}; x_{t_0}); x_{t_1})$. Note that to calculate EV with initial preferences or CV with final preferences, we must be able to separate taste shocks from income effects.

Macro welfare changes For each alternative micro welfare measure there is a corresponding macro welfare measure. For example, the *macro compensating variation* with *initial preferences* is

$$CV^M(A_{t_0}, L_{t_0}, A_{t_1}, L_{t_1}; x_{t_0}) = \phi,$$

where ϕ solves

$$V(A_{t_0}, L_{t_0}; x_{t_0}) = V(A_{t_1}, e^{-\phi} L_{t_1}; x_{t_0}).$$

In words, CV^M is the proportional change in final factor endowments necessary to make the planner with preferences $\succeq_{x_{t_0}}$ indifferent between the initial PPF (A_{t_0}, L_{t_0}) and PPF defined by $(A_{t_1}, e^{-\phi} L_{t_1})$.

Equation (14) in Proposition 5 applies using $\lambda^{cv}(A, L)$, the sales shares in a fictional economy with the PPF A, L but where consumers have stable homothetic preferences represented by the expenditure function $e^{cv}(p, u) = e(p, v_{t_0}, x_{t_0})u$ where $v_{t_0} = v(p_{t_0}, I_{t_0}; x_{t_0})$. Growth accounting for welfare is based on hypothetical sales shares evaluated at current technology but for fixed initial preferences and initial utility. The only information on preference parameters we need to know is elasticities of substitution at the final allocation. As discussed above, CV^M is equal to $-EV^M$ if we reverse the flow of time.

The gap between changes in welfare and real GDP is, to a second-order approximation (the analogue of that in Proposition 6) is

$$CV^M \approx \Delta \log Y - \frac{1}{2} \sum_{i \in N} \left[\Delta \log x' \frac{\partial \lambda_i}{\partial \log x} + \Delta \log V \frac{\partial \lambda_i}{\partial \log u} \right] \Delta \log A_i. \tag{41}$$

Proposition 7 can be also used to compute CV^M (instead of EV^M). To do this, we would still solve the differential equations for $d \log p^{ev}$ and $d\Psi^{ev}$, however, we would use different boundary conditions. The boundary conditions for compensating variation at initial preferences would match the data at t_0 instead of t_1, setting $p^{ev}_{t_0} = \mathbf{1}$ and $\Psi^{ev}_{t_0} = \Psi_{t_0}$. We would then solve the differential equations forward from t_0 to t_1 and use the resulting welfare-relevant shares to weight the technology shocks. This procedure effectively makes use of the fact that compensating variation at initial preferences going from t_0 to t_1 is equal to minus one times equivalent variation at final preferences if we go from t_1 to t_0. As with EV^M, conditional on the boundary conditions, we do not need to know the income elasticities or the taste shocks.

Appendix D Relation to Konüs Price Indices

A Konüs price index is defined as the ratio of the expenditure function at two different price systems holding fixed utility and preferences:

$$\frac{P_{t_1}(u, x)}{P_{t_0}(u, x)} = \frac{e(p_{t_1}, u; x)}{e(p_{t_0}, u; x)}.$$

Lemma 1 shows that EV^m can be calculated by deflating nominal income changes by the Konüs price index corresponding to final preferences and final utility (i.e. the final indifference curve).[46]

In the index number theory literature, it is common to work with Konüs price indices for some intermediate preferences or utility levels. For example, Diewert (1976), Caves et al. (1982), and Feenstra and Reinsdorf (2007). The advantage of this approach is that it requires far less information. For example, Diewert (1976) shows that a Tornqvist index of t_0 and t_1 measures the Konüs price index for a consumer with stable translog preferences with utility level $(u_{t_0} u_{t_1})^{\frac{1}{2}}$; Caves et al. (1982) and Feenstra and Reinsdorf (2007) prove a similar result for homothetic but unstable CES or translog preferences. In contrast to EV^m, these indices can all be computed *without* knowledge of any elasticities. In particular, these papers show that, under some assumptions (translog or CES), commonly used indices like

[46] Appendix C shows that compensating variation at initial preferences, CV^m, can be calculated by deflating nominal income changes by the Konüs price index corresponding to the *initial* indifference curve.

Tornqvist and Sato-Vartia do answer an economically meaningful question.

However, whilst these papers provide an interpretation for these commonly used indices, these indices do not measure EV or CV, which are of interest per se in applied micro and macro welfare analysis. Furthermore, these indices are not money metrics, as we show below. Our contribution, relative to common practice in the index number theory literature, is to instead characterize and analyze EV (and in Remark 2, Application 5.1, and Appendix C, CV). Furthermore, we provide a unified analysis of non-homotheticity and taste shocks, whereas the literature has tended to focus on one at a time under parametric assumptions or second-order approximations. We also compare partial versus general equilibrium notions of welfare.

To relate the aforementioned results to ours, consider the economic question that changes in nominal income between t_1 and t_0 deflated by a Konüs price index evaluated at some intermediate level of utility answers. For any base period t_b (which does not need to lie between t_0 and t_1) we can write

$$\log \frac{I_{t_1}}{I_{t_0}} - \log \frac{P_{t_1}(u_b, x_b)}{P_{t_0}(u_b, x_b)} = \left(\log \frac{I_{t_b}}{I_{t_0}} - \log \frac{P_{t_b}(u_b, x_b)}{P_{t_0}(u_b, x_b)} \right) + \left(\log \frac{I_{t_1}}{I_{t_b}} - \log \frac{P_{t_1}(u_b, x_b)}{P_{t_b}(u_b, x_b)} \right),$$

or

$$\log \frac{I_{t_1}}{I_{t_0}} - \log \frac{P_{t_1}(u_b, x_b)}{P_{t_0}(u_b, x_b)} = \log \frac{e(p_{t_0}, u_{t_b}; x_{t_b})}{e(p_{t_0}, u_{t_b}; x_{t_b})} - \log \frac{e(p_{t_1}, u_{t_b}; x_{t_b})}{e(p_{t_1}, u_{t_1}; x_{t_b})}.$$

The first summand on the right-hand side is $EV^m(p_{t_0}, I_{t_0}, p_{t_b}, I_{t_b}; x_b)$ and the second summand is $-EV^m(p_{t_1}, I_{t_1}, p_{t_b}, I_{t_b}; x_b)$.[47] In words, $\log \frac{I_{t_1}}{I_{t_0}} - \log \frac{P_{t_1}(u_b,x_b)}{P_{t_0}(u_b,x_b)}$ answers the question "For a consumer with preferences $\succeq_{x_{t_b}}$, what is the change in the t_0 endowment that makes her indifferent between her choice set at t_0 and t_b minus the change in the t_1 endowment that makes her indifferent between her choice set at t_1 and t_b?" In particular, note that the first term is in units of t_0 prices whereas the second one is in units of t_1 prices. Therefore, this is not a money metric that can be used for policy or counterfactual analysis. In sum, although our approach has stronger information requirements, it characterizes a widely-studied and fundamentally different object (i.e. a money metric) than what has commonly been studied in the index number theory literature.

Deaton and Muellbauer (1980) write in reference to the Konüs at intermediate utility result:

If we were willing to accept the reference indifference curve labelled by u^* (note: the

[47]Since CV^m at initial preferences can be calculated by reversing the flow of time and computing EV^m with the negative sign, $\log \frac{I_{t_1}}{I_{t_0}} - \log \frac{P_{t_1}(u_b,x_b)}{P_{t_0}(u_b,x_b)}$ can also be expressed as $-CV^m(p_{t_b}, I_{t_b}, p_{t_0}, I_{t_0}; x_b) + CV^m(p_{t_b}, I_{t_b}, p_{t_1}, I_{t_1}; x_b)$.

geometric average of a u_{t_0} and u_{t_1}) as the relevant one, this property of the Tornqvist index is attractive since the quadratic specification can provide a second-order approximation to any arbitrary cost function. Without knowing the parameters of the cost function, we lack more specific information about the reference indifferent curve (such as what budget level and price vector correspond to it), and the result is of no help in constructing a constant utility cost-of-living index series with more than two elements [elements refer to time periods]. A chained series of pairwise Tornqvist indices can always be constructed, but this has a different reference indifference curve for every link in the chain. (Deaton and Muellbauer, 1980, page 174)

That is, in practice most index numbers are constructed by chaining, but the intermediate utility result does not apply to chained indices unless the path of prices is linear (Feenstra and Reinsdorf, 2000). In our paper we characterize how equivalent variation at final preferences or compensating variation at initial preferences differ from chained (Divisia) indices under arbitrary price and income paths.

Appendix E Non-homothetic CES preferences

In this appendix, we derive (7). We also compare EV^m with the utility index (under a popular cardinalization) in non-homothetic CES preferences and show that changes in the utility index are not equal to changes in equivalent or compensating variation.

E.1 Derivation of Marshallian budget shares

This appendix provides a derivation of the log-linearized expression (7). Changes in Marshallian budget share (before imposing functional forms and suppressing dependence on t) are given by

$$\begin{aligned} d\log b_i &= d\log p_i - d\log I + \sum_j \varepsilon_{ij}^M d\log p_j + \varepsilon_i^w d\log I + d\log x_i, \\ &= d\log p_i - d\log I + \sum_j \left(\varepsilon_{ij}^H - \varepsilon_i^w b_j\right) d\log p_j + \varepsilon_i^w d\log I + d\log x_i, \end{aligned}$$

where ε^H and ε^M are the Hicksian and Marshallian price elasticities, $\varepsilon^w$are the income elasticities, and $d\log x_i$ is a residual that captures changes in shares not attributed to changes in prices or income. The second equality is an application of Slutsky's equation. Since b_i are expenditure shares that always add up to one, it must be that $\sum_i d\log x_i = 0$ and $\sum_i b_i\varepsilon_i^w = 1$.

When preferences are non-homothetic CES, the expenditure function can be written as (using a popular cardinalization)

$$e(p,u) = \left(\sum_i \omega_i p_i^{1-\theta_0} u^{\xi_i} \right)^{\frac{1}{1-\theta_0}}, \tag{42}$$

and Hicksian demand as

$$c_i = \omega_i \left(\frac{p_i}{\sum_j p_j c_j} \right)^{-\theta_0} u^{\xi_i}, \tag{43}$$

where ω_i and ξ_i are some parameters. The Hicksian price elasticity for $j \neq i$ is

$$\frac{\partial \log c_i}{\partial \log p_j} = \varepsilon_{ij}^H = \theta_0 \frac{p_j c_j}{I} = \theta_0 b_j.$$

Using this fact and the identity $\varepsilon_{ii}^H = -\sum_{j\neq i} \varepsilon_{ij}^H$, we can rewrite changes in budget shares as

$$\begin{aligned}
d\log b_i &= \sum_j \left(\varepsilon_{ij}^H - \varepsilon_i^w b_j \right) d\log p_j + d\log p_i + (\varepsilon_i^w - 1)\, d\log I + d\log x_i \\
&= \left(1 - \sum_{j\neq i} \varepsilon_{ij}^H \right) d\log \frac{p_i}{I} + \sum_{j\neq i} \varepsilon_{ij}^H d\log \frac{p_j}{I} + \varepsilon_i^w \left[d\log I - \sum_j b_j d\log p_j \right] + d\log x_i \\
&= \left(1 - \sum_{j\neq i} \theta_0 b_j \right) d\log \frac{p_i}{I} + \sum_{j\neq i} \theta_0 b_j d\log \frac{p_j}{I} + \varepsilon_i^w \left[d\log I - \sum_j b_j d\log p_j \right] + d\log x_i \\
&= (1 - \theta_0(1-b_i))\, d\log \frac{p_i}{I} + \sum_{j\neq i} \theta_0 b_j d\log \frac{p_j}{I} + \varepsilon_i^w \left[d\log I - \sum_j b_j d\log p_j \right] + d\log x_i \\
&= (1-\theta_0) \left[d\log p_i - \sum_j b_j d\log p_j \right] + (\varepsilon_i^w - 1) \left[d\log I - \sum_j b_j d\log p_j \right] + d\log x_i.
\end{aligned} \tag{44}$$

In the body of the paper we use ε_i in place of ε_i^w and we explicitly index by t. This completes the derivation of (7) where $d\log x$ could be any perturbation to budget shares not explained by income and substitution effects.

For ex-ante counterfactuals, we can use (44) as a differential equation to solve for budget shares in the future. To do this, we must put some structure on $d\log x$. For example, assume that $d\log x$ is being driven by changes in taste parameters ω_i, and θ_0 and ξ_i are constant.

Then, differentiating (42) and (43) at any point t,

$$d \log b_{it} = d \log \omega_{it} + (1 - \theta_0) \left(d \log p_{it} - d \log I_t \right) + \xi_i d \log u_t, \tag{45}$$

and

$$d \log u_t = \frac{1-\theta}{\sum_j b_{jt} \xi_j} \left[d \log I_t - \sum_j b_{jt} d \log p_{jt} \right] - \frac{1}{\sum_j b_{jt} \xi_j} \sum_j b_{jt} d \log \omega_{jt}. \tag{46}$$

Substituting (46) into (45), equations (7) and (44) can be written at any point t as

$$d \log b_{it} = (1 - \theta_0) \left[d \log p_{it} - \mathbb{E}_{b_t}(d \log p_t) \right] + (\varepsilon_{it}^w - 1) \left[d \log I_t - \mathbb{E}_{b_t}(d \log p_t) \right] + d \log x_{it},$$

with demand shocks

$$d \log x_{it} = d \log \omega_{it} - \frac{\xi_i}{\sum_j b_{jt} \xi_j} \sum_j b_{jt} d \log \omega_{jt}, \tag{47}$$

and income elasticities

$$\varepsilon_{it}^w - 1 = (1 - \theta_0) \left(\frac{\xi_i}{\sum_j b_{jt} \xi_j} - 1 \right). \tag{48}$$

This is a differential equation that pins down budget shares b as a function of prices, incomes and primitives ω. Hence, it can be solved numerically for any path of prices, incomes, and ω to arrive at b.

Note that $\sum_i b_i \xi_i$ is not pinned down by observables since scaling all ξ proportionally does not change any observable.

E.2 Comparison of welfare and changes in utility index

In this appendix, we discuss the difference between changes in welfare as measured by the equivalent variation and changes in the utility index in non-homothetic CES preferences. This utility index is used in Section IIIA of Redding and Weinstein (2020) as a welfare measure. We show that there is no normalization of the parameters such that the equivalent variation (or the compensating variation) is equal to changes in the utility index unless preferences are homothetic and stable.

In this section, for brevity we assume away taste shocks (for taste shocks, see Appendix F). The micro equivalent variation is given by

$$EV^m = \log \frac{e(p_{t_0}, v(p_{t_1}, I_{t_1}))}{e(p_{t_0}, v(p_{t_0}, I_{t_0}))},$$

where $v(p, I)$ is the indirect utility function, initial prices and income are p_{t_0} and I_{t_0}, and final prices and income are p_{t_1} and I_{t_1}.

The utility index u at t is equal to $v(p_t, I_t)$, and can be calculated by solving for u in $I_t = e(p_t, u)$. Equivalently, one can calculate changes in u_t using the price index $P_t \equiv e(p_t, u_t)/u_t$. The change in the utility index between t_0 and t_1 is given by

$$U \equiv \log \frac{v(p_{t_1}, I_{t_1})}{v(p_{t_0}, I_{t_0})}.$$

As this definition makes clear, EV and U are not generically the same. In particular, whereas EV can be defined in terms of a hypothetical choice and is independent of the utility function chosen to represent preferences (how much income would the household need to be given to make them indifferent), U will depend on the cardinal properties of the utility function.

Consider the expenditure function in equation (42). If preferences are homothetic ($\xi_i = \bar{\xi}$ for all i), then $e(p, u) = \left(\sum_i \omega_i p_i^{1-\theta_0}\right)^{\frac{1}{1-\theta_0}} u^{\frac{\bar{\xi}}{1-\theta_0}}$ and we can write

$$EV^m = \frac{\bar{\xi}}{1-\theta_0} U.$$

So, when preferences are homothetic, in order for $EV^m = U$ we must cardinalize utility by setting $\bar{\xi} = 1 - \theta_0$ so that the expenditure function is homogeneous of degree 1 in u ($d \log e / d \log u = 1$). In other words, although there are infinitely many utility functions that represent these preferences, when preferences are homothetic, there is one representation where $EV^m = U$.

We now consider the non-homothetic case, and we characterize the difference between EV^m and U to a first and second order. We write these results in terms of primitive shocks (that is, changes in income and prices) rather than in terms of changes in endogenous objects like budget shares.

Using Proposition 2, we have that to a first-order EV^m is

$$dEV^m = d \log e - b d \log p = d \log Y,$$

where $d \log Y$ is the first-order change in real consumption as measured by Tornqvist or Divisa (to a first-order, they are equivalent). Hence, to a first order, Tornqvist and EV are the same. The second-order change in EV^m is, by Proposition 2, equal to

$$\begin{aligned} d^2 EV^m &= d^2 \log e - dbd \log p - (d \log e - bd \log p)\, Cov_b(\varepsilon^w, d \log p) \\ &= d^2 \log Y - (d \log e - bd \log p)\, Cov_b(\varepsilon^w, d \log p), \end{aligned}$$

where ε^w is the vector of income elasticities and $d^2 \log Y$ is the change in real consumption as measured by a Tornqvist or Divisa index (to a second-order, they are equivalent). On the other hand, the first and second-order changes in the utility index are given by (derivations are available upon request)

$$dU = \frac{1-\theta_0}{\sum_i b_i \bar{\xi}} \left(d \log e - bd \log p \right),$$

$$\begin{aligned} d^2 U = \frac{1-\theta_0}{\sum_i b_i \xi_i} \Bigg[& d^2 \log e - dbd \log p - (d \log e - bd \log p) \sum_i b_i (\varepsilon_i^w - 1) d \log p_i \\ & - \frac{1}{1-\theta_0} \sum_i b_i \varepsilon_i^w \left((\varepsilon_i^w - 1) \right) \left(d \log e - bd \log p \right)^2 \Bigg] \end{aligned}$$

The derivatives EV^m and U are in general different. Whereas EV^m is a function of observables, U depends on the cardinalization of the utility function. In particular $\sum_i b_i \xi_i$ affects the response of U but is not a primitive parameter of the ordinal preference relation, and hence is not pinned down by observables, as discussed in Section E.1. A standard approach in the literature to pin down $\sum_i b_i \xi_i$ is to set one of the ξ to 1.

Now we compare the first and second-order derivatives in turn. The first order difference is

$$dU - dEV^m = \left(\frac{1-\theta_0}{\sum_i b_i \xi_i} - 1 \right) \left(d \log e - bd \log p \right).$$

If we impose a normalization on utility parameters such that, in the initial point,

$$\frac{1-\theta_0}{\sum_i b_i \xi_i} = 1,$$

we have that $dU = dEV^m = d \log Y$. This normalization is effectively ensuring that $\partial \log e / \partial \log u = 1$.

Now let's consider the second-order difference and let's impose the same normalization

$$
\begin{aligned}
d^2U - d^2EV^m &= -\frac{1}{1-\theta_0}\sum_i b_i\varepsilon_i^w(\varepsilon_i^w - 1)\,(d\log e - bd\log p)^2 \\
&\quad -(d\log e - bd\log p)\left[\sum_i b_i(\varepsilon_i^w - 1)\sum_i b_i d\log p_i\right] \\
&= -\frac{1}{1-\theta_0}\sum_i b_i\varepsilon_i^w(\varepsilon_i^w - 1)\,(d\log e - bd\log p)^2
\end{aligned}
$$

$$
= -\frac{1}{1-\theta_0}Var_b(\varepsilon_i^w)\,(d\log e - bd\log p)^2 \neq 0,
$$

where we used $\sum_i b_i\varepsilon_i^w = 1$. Hence, unless preferences are homothetic (in which case $\varepsilon_i^w = 1$ for every i), the change in U and EV^m are not the same even under the normalization. This is not to mention that globally, we cannot ensure that the normalization

$$
\frac{1-\theta_0}{\sum_i b_i\xi_i} = 1
$$

always holds. This means that the gap between EV^m and U, which exists at the initial equilibrium, only gets more severe if, once we commit to a specific normalization of utility, $\frac{1-\theta_0}{\sum_i b_i\xi_i}$ starts to change from 1.

Recall from Appendix C that changes in real consumption are equal to an average of equivalent and compensating variation, up to a second order approximation. Since changes in the utility index are not equal to a Tornqvist real consumption index, it follows that the utility index is not equal to an average of EV and CV.

Appendix F Comparison of Quality and Taste Changes

In this appendix, we discuss how our welfare results can be extended to environments with unobserved quality changes. In this appendix, we also contrast the bias we identify with the "taste shock bias" discussed by Redding and Weinstein (2020).

As mentioned in Footnote 9, the standard approach to measuring quality change is hedonics. For example, suppose that consumers have CES preferences (indexed by tastes x) over q_ic_i across different varieties i, where c_i is the quantity and q_i is the quality of good i. For example, each i is a different variety of chocolate, c_i is the number of boxes of chocolate, and q_i is the weight of each box of chocolate i. So the characteristic that consumers have preferences over is the total weight of chocolate they purchase of each type, and consumers do not care about how many boxes their chocolate came in.

Under these assumptions, quality changes are equivalent to changes in prices, so we can write the quality-adjusted price of good i as $p_i = \tilde{p}_i / q_i$, where $\tilde{p}_i$ is the observed market price of good i. In our example, $\tilde{p}_i$ is the observed price per box and $\tilde{p}_i / q_i$ is the price per ounce. Changes in quality-adjusted prices are given by $\Delta \log p_i = \Delta \log \tilde{p}_i - \Delta \log q_i$. On the other hand, changes in x indicate changes in preferences for the different varieties of chocolate. Unlike changes in q_i, which are measured in ounces per box, changes in x do not have interpretable units and the effects on the utility index depend on the choice of cardinalization.

Substituting this into our various propositions allows us to isolate the way quality changes affect our results and how they compare with changes in tastes. For example, Proposition 2 becomes the following (for brevity, we assume homothetic CES preferences):

Proposition 12 (Approximate Micro with Quality Change)**.** *Consider some perturbation in demand $\Delta \log x$, market prices $\Delta \log \tilde{p}$, quality $\Delta \log q$, and income $\Delta \log I$. Then, to a second-order approximation, the change in real consumption is*

$$\begin{aligned}\Delta \log Y \approx \Delta \log I - \mathbb{E}_b \left(\Delta \log \tilde{p} \right) - \frac{1}{2}(1-\theta_0) Var_b(d \log \tilde{p}) \\ + \frac{1}{2}(1-\theta_0) Cov_b(d \log q, d \log \tilde{p}) - \frac{1}{2} Cov_b(d \log x, d \log \tilde{p}),\end{aligned}$$

and the change in welfare is

$$\begin{aligned}EV^m \approx \Delta \log I - \mathbb{E}_b \left(\Delta \log \tilde{p} - \Delta \log q \right) - \frac{1}{2}(1-\theta_0) Var_b \left(\Delta \log \tilde{p} \right) \\ - \frac{1}{2}(1-\theta_0) Var_b \left(\Delta \log q \right) + (1-\theta_0) Cov_b \left(\Delta \log \tilde{p}, \Delta \log q \right) - Cov_b \left(\Delta \log x, \Delta \log p \right),\end{aligned}$$

where $\mathbb{E}_b(\cdot)$, $Var_b(\cdot)$, and $Cov_b(\cdot)$ are evaluated using budget shares at t_0 as probability weights.

Hence, by subtracting these two expressions, we can derive the gap between real consumption and welfare up to a second order approximation as

$$\begin{aligned}EV^m - \Delta \log Y \approx \underbrace{\mathbb{E}_b \left(\Delta \log q \right)}_{\text{average quality}} + \frac{1}{2} \underbrace{(\theta_0 - 1) Var_b \left(\Delta \log q \right)}_{\text{dispersion in quality}} + \frac{1}{2} \underbrace{(1-\theta_0) Cov_b \left(\Delta \log \tilde{p}, \Delta \log q \right)}_{\text{covariance of price and quality}} \\ - \frac{1}{2} \underbrace{Cov_b \left(\Delta \log x, \Delta \log \tilde{p} \right)}_{\text{covariance of taste and price}} + \underbrace{Cov_b \left(\Delta \log x, \Delta \log q \right)}_{\text{covariance of taste and quality}} . \quad (49)\end{aligned}$$

The first term on the right-hand side captures how the average increase in quality raises welfare relative to real consumption. The second term captures the fact that dispersion in quality raises welfare if the elasticity of substitution is greater than one (since the consumer substitutes towards goods with relatively higher quality, but quality is not captured by market prices in real consumption). The third term is an interaction (cross-partial) effect that raises welfare if market prices fall for goods whose quality rose, as long as the elasticity of substitution is greater than one. The fourth term is the bias we have been emphasizing in the paper so far. The final term is the interaction between quality and taste changes — welfare is higher, at final preferences, if tastes increase for goods whose quality also increase.

In our analysis, we assume that prices have already been adjusted for quality so the only non-zero term is the fourth one. In other words, in the body of the paper, we assume that $\Delta \log q = 0$, which means that (49) simplifies to

$$EV^m - \Delta \log Y \approx -\frac{1}{2} Cov_b \left(\Delta \log x, \Delta \log \tilde{p}\right). \tag{50}$$

Welfare is higher than real consumption if the covariance between taste shocks and prices is negative. This is independent of the value of the elasticity of substitution.

Comparison to Redding and Weinstein (2020). We can use (49) to contrast our approach to that of Redding and Weinstein (2020). The "taste shifters" in that paper are mathematically equivalent to quality shocks ($\Delta \log q \neq 0$), and preferences are stable over "taste-adjusted consumption" ($\Delta \log x = 0$). Equation (49) simplifies to

$$EV^m - \Delta \log Y \approx \mathbb{E}_b \left(\Delta \log q\right) + \frac{1}{2}(\theta_0 - 1) Var_b \left(\Delta \log q\right) - \frac{1}{2}(\theta_0 - 1) Cov_b \left(\Delta \log \tilde{p}, \Delta \log q\right). \tag{51}$$

Comparing (50) to (51) elucidates the differences. First, the average level of $\Delta \log q$ affects welfare but the average level of $\Delta \log x$ does not. Redding and Weinstein (2020) assume that unweighted average of $\Delta \log q$ is zero.[48] Second, for shocks to $\Delta \log q$, even when they are mean zero, dispersion in q can raise or lower welfare depending on the elasticity of substitution. Hence, shocks to q on their own can change welfare, holding prices and income constant, and the sign of this effect depends on the elasticity of substitution. This is in contrast to shocks to x which cannot change money-metric welfare on their own if prices and income are held constant. Third, in both (50) and (51), the covariance of taste shifters and market prices matters, however, in (51) the sign of the covariance depends on whether the elasticity of substitution is greater than or less than one, whereas in (50), the

[48] As discussed earlier, if $\Delta \log q$ is interpreted as a taste shock rather than a quality shock, then there is nothing in the data that pins down the average level of $\Delta \log q$ since it is not a primitive of the ordinal preference relation.

sign is always the same.

Appendix G Analytical Examples with Input-Output Connections

We first discuss some differences between Proposition 2 and Proposition 6 in the presence of intermediate inputs. Proposition 2 shows that if all price changes are the same, there can be no gap between micro welfare EV^m and real consumption. The general equilibrium counterpart of this statement is not true. That is, there can be a gap between real GDP and welfare even if all productivity shocks are the same. Specifically, suppose that productivity growth is common across all goods ($\Delta \log A_i = \Delta \log A > 0$) and denote the gross output to GDP ratio by $\lambda^{sum} = \sum_{i \in N} \lambda_i \geq 1$. Then Proposition 6 implies that the gap between real GDP and welfare is

$$EV^M - \Delta \log Y \approx \frac{1}{2} \left[\Delta \log x' \frac{\partial \lambda^{sum}}{\partial \log x} + \Delta \log V \frac{\partial \lambda^{sum}}{\partial \log u} \right] \Delta \log A, \tag{52}$$

where the term in square brackets is the change in the gross-output-to-GDP ratio due to demand-side forces only. In particular, if demand shifts towards sectors with higher value-added as a share of sales, then $EV^M < \Delta \log Y$. Intuitively, this happens because welfare is less reliant on intermediates than real GDP, and hence real GDP is more sensitive to productivity shocks. Of course, in the absence of intermediate inputs, this effect disappears because λ^{sum} will always equal one.

In our quantitative results in Application 1 (section 5.1), the reallocation in sales towards sectors with lower intermediate input use accounts for roughly 18% of the gap between constant-initial-sales shares TFP and aggregate TFP growth, and 35% of the gap between aggregate TFP growth and welfare-relevant TFP growth.

We now extend the analytic examples in Section 4.2 to show how input-output connections can amplify or mitigate the gap between macro welfare EV^M and real GDP $\Delta \log Y$. For models with linear PPFs, input-output connections affect the gap between real GDP and welfare in two ways: (1) the impact of technology shocks is bigger when there are input-output linkages because $\Psi \geq Id$ and $\lambda_i \geq b_i$; (2) the production network "mixes" the shocks, and this may reduce the correlation of supply and demand shocks by making the technology shocks more uniform. However, since it is the covariance (not the correlation) of the shocks that matters, this means the effects are, at least theoretically, ambiguous.

To see these two forces, consider the three economies depicted in Figure 4. Each of these economies has a roundabout structure. Panel 4a depicts a situation where each producer uses only its own output as an input, Panel 4b a situation where all producers use the same basket of goods (denoted by M) as an intermediate input, and Panel 4c a situation where each producer uses the output of the other producer as an input. We compute the correction to GDP necessary to arrive at welfare for each of these cases using Proposition 8. For clarity, we focus on demand shocks caused by instability rather than non-homotheticity, though it should be clear that this does not affect any of the intuitions.

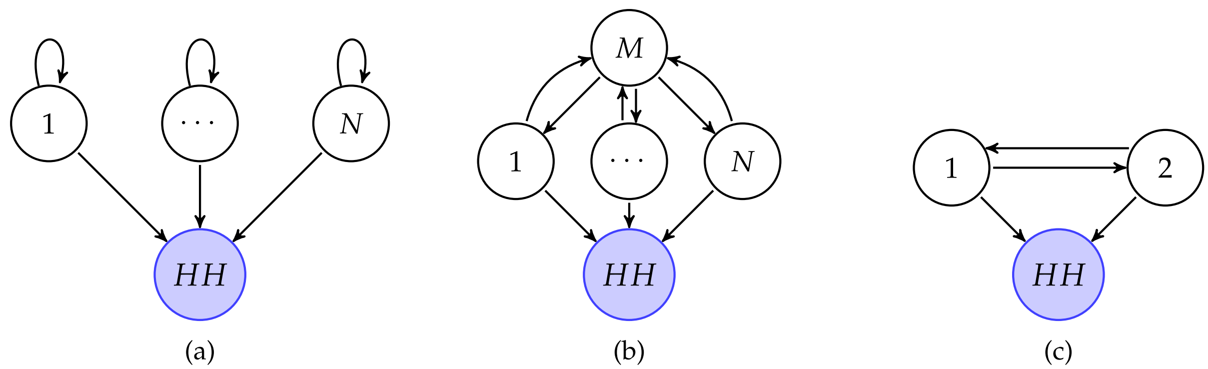


Figure 4: Three different kinds of round-about economy. The arrows represent the flow of goods. The only factor is labor which is not depicted in the diagram.

For Panel 4a, we get

$$EV^M - \Delta \log Y \approx \frac{1}{2} Cov_b(\Delta \log x_i, \Omega_{iL}^{-1} \Delta \log A_i),$$

where the covariance is computed across goods $i \in N$ and Ω_{iL} is the labor share for i. Hence, as intermediate inputs become more important, the necessary adjustment becomes larger. This is because, for a given vector of preference shocks, the movement in sales shares is now larger due to the roundabout nature of production.[49]

On the other hand, for Panel 4b, we get[50]

$$EV^M - \Delta \log Y \approx \frac{1}{2} \left(Cov_b(\Delta \log x_i, \Delta \log A_i) - Cov_b(\Delta \log x_i, \Omega_{iL}) \frac{\sum_{i \in N} \Delta \log A_i}{\sum_{i \in N} \Omega_{iL}} \right).$$

Hence, in this case, if the labor share Ω_{iL} is the same for all $i \in N$, then the intermediate input share is irrelevant. Intuitively, in this case, all producers buy the same share of materials, so a shock to the composition of household demand does not alter the sales of any

[49]Even if all productivity shocks are the same, there may still be an adjustment due to heterogeneity in labor shares. In particular, if demand shocks are higher for sectors with higher labor shares, then $EV^M < \Delta \log Y$ when technology shocks are positive.

[50]For this example, we assume that there are no productivity shocks to the intermediate bundle $\Delta \log A_M = 0$ and we assume that $\Omega_{iM} = 1/N$ for each $i \in N$.

producer through the supply chain, and hence only the first-round non-network component of the shocks matters.[51]

Finally, consider Panel 4c. For clarity, we focus on the case where only producer 1 gets a productivity shock ($\Delta \log A_2 = 0$). In this case, the difference between real GDP and welfare is

$$EV^M - \Delta \log Y \approx \frac{1}{2}\frac{1}{1-\Omega_{12}\Omega_{21}} Cov_b \left(\Delta \log x, \begin{bmatrix} 1 \\ \Omega_{21} \end{bmatrix} \right) \Delta \log A_1.$$

As the intermediate input share Ω_{21} approaches one, the adjustment goes to zero (since the covariance term goes to zero). Intuitively, as Ω_{21} goes to one, the increase in demand for the first producer from a change in preferences is exactly offset by a reduction in demand from the second producer who buys inputs from the first producer. In this limiting case, changes in consumer preferences have no effect on the overall sales share of the first producer.

To recap, in the first, second, and third example the gap between welfare and real consumption increases, is independent of, and decreases in the intermediate input share. Hence, the effect of input-output networks on the adjustment are potent but theoretically ambiguous.

Appendix H Additional details on the Baumol application

In this appendix, we provide some intuition for why, from a welfare perspective, structural transformation caused by income effects or taste shocks is roughly twice as important as that caused by substitution effects. We also use a structural nested-CES model to explore the change in welfare-relevant TFP outside of the two polar extremes in Section 5.

H.1 Intuition for size of welfare-adjustment

According to our results in Section 5, structural transformation caused by income effects or demand instability reduced welfare by roughly twice as much as structural transformation caused by substitution effects. To understand why the necessary adjustment is roughly

[51] As indicated in Footnote 49, if the labor share is heterogeneous across producers, there is an additional adjustment which depends on the covariance between demand shocks and labor shares. If the demand shocks reallocate expenditures towards sectors with high labor shares, then welfare becomes less sensitive to productivity shocks than real GDP.

twice as big, consider the second-order approximation in Proposition 6:

$$\Delta \log TFP^{\text{welfare}} \approx \Delta \log TFP + \frac{1}{2} \sum_{i \in N} \left[\Delta \log x \frac{\partial \lambda_i}{\partial \log x} + \Delta \log V \frac{\partial \lambda_i}{\partial \log u} \right] \Delta \log A_i, \tag{53}$$

where

$$\Delta \log TFP \approx \sum_{i \in N} \lambda_{it_0} \Delta \log A_i + \frac{1}{2} \sum_{i \in N} \Delta \lambda_i \Delta \log A_i.$$

If changes in sales shares are due entirely to demand-driven factors, then the term in square brackets in (53) is equal to $\Delta \lambda_i$, so

$$\Delta \log TFP^{\text{welfare}} \approx \sum_{i \in N} \lambda_{it_0} \Delta \log A_i + \sum_{i \in N} \Delta \lambda_i \Delta \log A_i.$$

In other words, the adjustment to the initial sales shares must be roughly twice as large as the adjustment to the initial sales shares caused by substitution effects.[52]

H.2 Welfare-TFP outside of polar extremes

In practice, both substitution effects and non-homotheticities are likely to play an important role in explaining structural transformation. To dig deeper into the size of the welfare adjustment outside our two polar cases, we use a simplified version of the model introduced in Section 4 calibrated to the US economy, accounting for input-output linkages and complementarities, and use the model to quantify the size of the welfare-adjustment as a function of the elasticities of substitution. We calculate TFP by industry in the data allowing for cross-industry variation in capital and labor shares. For simplicity, we feed these TFP shocks as primitive shocks into a one-factor model.

Remarkably, Proposition 5 implies that to compute the welfare-relevant change in TFP, we must only supply the information necessary to compute λ^{ev}. That is, since we know sales shares in the terminal period 2014, we do not need to model the non-homotheticities or demand-shocks themselves, and the exercise requires no information on the functional form of non-homotheticities or the slope of Engel curves or magnitude of income elasticities *conditional* on knowing the elasticities of substitution.

We map the model to the data as follows. We assume that the constant-utility final de-

[52]These second-order approximations are more accurate if changes in sales shares are well-approximated by linear time trends, and the surprising accuracy of the second-order approximation is a result of this fact.

mand aggregator has a nested-CES form. There is an elasticity θ_0 across the three groups of industries: primary, manufacturing, and service industries. The inner nest has elasticity of substitution θ_1 across industries within primary (2 industries), manufacturing (24 industries), and services (35 industries).[53] Production functions are also assumed to have nested-CES forms: there is an elasticity of substitution θ_2 between the bundle of intermediates and value-added, and an elasticity of substitution θ_3 across different types of intermediate inputs. For simplicity, we assume there is only one primary factor of production (a composite of capital and labor). We solve the non-linear model by repeated application of Proposition 7 in the fictional economy with stable and homothetic preferences.

We calibrate the CES share parameters so that the model matches the 2014 input-output tables provided by the BEA. For different values of the elasticities of substitution $(\theta_0, \theta_1, \theta_2, \theta_3)$ we feed changes in industry-level TFP (going backwards, from 2014 to 1947) into the model and compute the resulting change in aggregate TFP. This number represents the welfare-relevant change in aggregate TFP. We report the results in Table 2.

Table 2: Percentage change in measured and welfare-relevant TFP in the US from 1947 to 2014.

$(\theta_0, \theta_1, \theta_2, \theta_3)$	(1,1,1,1)	(0.5,1,1,1)	(1,0.5,1,1)	(1,1,0.5,1)	(1,1,1,0.5)
Welfare TFP	46%	46%	54%	48%	55%
Measured TFP	60%	60%	60%	60%	60%

The first column in Table 2 shows the change in welfare-relevant TFP assuming that there are no substitution effects (all production and consumption functions are Cobb-Douglas). In this case, all changes in sales shares in the data are driven by non-homotheticities or demand-instability, and hence welfare-relevant TFP has grown more slowly than measured TFP, exactly as discussed in the previous section. The other columns show how the results change given lower elasticities of substitution. As we increase the strength of complementarities (so that substitution effects are active), the implied non-homotheticities required to match changes in sales shares in the data are weaker. This in turn reduces the gap between measured and welfare-relevant productivity growth.

Table 2 also shows that not all elasticities of substitution are equally important. The results are much more sensitive to changes in the elasticity of substitution across more disaggregated categories, like materials, than aggregated categories, like agriculture, manufacturing, and services.

[53]In order to map this nested structure to our baseline model, good 0 is a composite of good 1-3, where good 1 is a composite of primary industries, good 2 is a composite of manufacturing industries, and good 3 is a composite of service industries. Goods 4-65 are the disaggregated industries. Finally, good 66 is the single factor of production.

To see why the results in Table 2 are differentially sensitive to changes in different elasticities of substitution, we use Proposition 8 to obtain the following second-order approximation:

$$\Delta \log TFP^{\text{welfare}} \approx \sum_i \lambda_i \Delta \log A_i + \frac{1}{2} \sum_{j \in \{0\}+\mathcal{N}} (\theta_j - 1) \lambda_j Var_{\Omega_{(j,:)}} \left(\sum_{k \in \mathcal{N}} \Psi_{(:,k)} \Delta \log A_k \right), \quad (54)$$

where λ, Ω, and Ψ are evaluated at t_1. The second term is half the sum of changes in Domar weights due to substitution effects (i.e. changes in welfare-relevant sales shares) times the change in productivities. Note that changes in these welfare-relevant sales shares are linear in the microeconomic elasticities of substitution. The importance of some elasticity θ depends on

$$\sum_j \lambda_j Var_{\Omega_{(j,:)}} \left(\sum_{k \in \mathcal{N}} \Psi_{(:,k)} \Delta \log A_k \right),$$

where the index j sums over all CES nests whose elasticity of substitution is equal to θ (i.e. all j such that $\theta_j = \theta$). Therefore, elasticities of substitution are relatively more potent if: (1) they control substitution over many nests with high sales shares λ_j, or (2) if the nests corresponding to those elasticities are heterogeneously exposed to the productivity shocks.

We compute the coefficients in (54) for our model's various elasticities using the IO table at the end of the sample. The coefficient on $(\theta_0 - 1)$, the elasticity of substitution between agriculture, manufacturing, and services in consumption is only 0.01. This explains why the results in Table 2 are not very sensitive to this elasticity. On the other hand, the coefficient on $(\theta_1 - 1)$, the elasticity across disaggregated consumption goods, is much higher at 0.21. The coefficient on $(\theta_2 - 1)$, the elasticity between materials and value-added bundles is 0.07. Finally, the coefficient on $(\theta_3 - 1)$, the elasticity between disaggregated categories of materials is 0.25. This underscores the fact that elasticities of substitution are more important if they control substitution in CES nests which are very heterogeneously exposed to productivity shocks — that is, nests that have more disaggregated inputs.

According to equation (54), setting $\theta_1 = \theta_2 = \theta_3 = 1$ (which is similar to abstracting from heterogeneity within the three broader sectors and heterogeneity within intermediate inputs), then θ_0 is the only parameter that can generate substitution effects in the model. This may help understand why more aggregated models of structural transformation (e.g. Buera et al., 2015 and Alder et al., 2019) require low values of θ_0 to account for the extent of sectoral reallocation in the data.

H.3 The Baumol effect in real consumption.

We now show that similar conclusions apply if we measure welfare changes using Personal Consumption Expenditures (PCE) data from the BEA on consumer prices and budget shares across 66 categories of goods and services in the US between 1947 and 2019.

Specifically, we measure the change in microeconomic welfare using Lemma 1 under alternative assumptions about income and substitution elasticities. We apply this formula statically and calculate welfare changes between two time periods taking as given changes in prices and nominal expenditures. Under the assumptions of Proposition 9, these static numbers also represent the change in macroeconomic welfare in a dynamic model between two steady-states.

If changes in budget shares are driven by substitution effects only, then welfare changes are equal to growth in real consumption per capita. If changes in budget shares are driven by income effects and demand instability only, then welfare changes between any year and 2019 are given by changes in nominal expenditures deflated by a price index using 2019 budget shares.

Figure 5 shows that for comparisons that are close to 2019, the change in welfare is not very sensitive to the assumptions on demand instability and income effects versus substitution effects because, at high frequency, the shocks are small and the sales shares at our level of aggregation (66 goods and services) are stable. On the other hand, for longer time periods, welfare growth is smaller if changes in budget shares took place due to income effects (or demand instability) rather than substitution effects. That is, comparing 1947 and 2019, the change in welfare per capita was 145 log points if preferences are homothetic and stable, but it was only 126 log points if changes in budget shares were entirely due to demand shocks and income effects. As before, structural transformation in consumption caused by demand shocks and income effects is roughly twice as important for welfare as structural transformation caused by substitution effects.

Appendix I Within-Industry Supply and Demand Shocks

In this appendix, we introduce a specification of our model with an explicit firm-industry structure. We show that within-industry supply and demand shocks can also drive a wedge between welfare and real GDP, and we show that this gap is linearly separable (to a second-order) from across-industry biases. For simplicity, we abstract from non-homotheticities.

Definition 7 (Industrial Structure)**.** An economy has an *industry structure* if the following

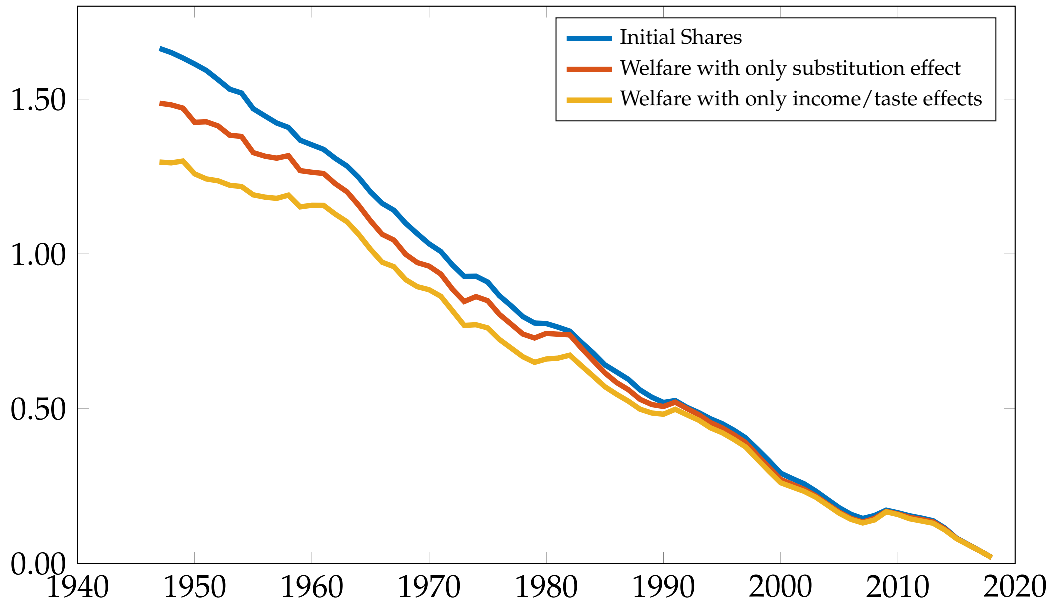


Figure 5: Change in welfare per capita from 1947 to 2019 using PCE prices and expenditures for 66 goods and services from the BEA. The blue line uses initial shares (in each year t between 1947 and 2019) to calculate the deflator. The red and yellow line measure the increase in welfare between t and 2019 under alternative assumptions about income and substitution elasticities. The red line assumes that budget shares change only due to substitution effects (welfare is equal to measured chained-real consumption). The yellow line assumes that budget shares change only due to income effects (or demand instability).

conditions hold:

i. Each firm i belongs to one, and only one, industry I. Firms in the same industry share the same constant-returns-to-scale production function up to a firm-specific Hicks-neutral productivity shifter A_i.

ii. The representative household has homothetic preferences over industry-level goods, where the Ith industry-level consumption aggregator is

$$c_I = \left(\sum_{i \in I} \bar{b}_{iI} x_i c_i^{\frac{\zeta_I - 1}{\zeta_I}} \right)^{\frac{\zeta_I}{\zeta_I - 1}},$$

where c_i are consumption goods purchased by the household from firm i in industry I and x_i are firm-level demand shocks.

iii. Inputs purchased by any firm j from firms i in industry I are aggregated according to

$$m_{jI} = \left(\sum_{i \in I} \bar{s}_{iI} m_{ji}^{\frac{\sigma_I - 1}{\sigma_I}} \right)^{\frac{\sigma_I}{\sigma_I - 1}},$$

where m_{ji} are inputs purchased by firm j from firm i, and $\bar{s}_{iI}$ is a constant.

Input-output and production network models that are disciplined by industry-level data typically have an industry structure of the form defined above. For such economies, the following proposition characterizes the bias in real GDP relative to welfare.

Proposition 13 (Aggregation Bias)**.** *For models with an industry structure, in response to firm-level supply shocks $\Delta \log A$ and demand shocks $\Delta \log x$, we have*

$$\Delta \log EV^M \approx \Delta \log Y + \frac{1}{2} \sum_I b_I Cov_{b_{(I)}}(\Delta \log x, \Delta \log A) + \Theta,$$

where b_I is industry I's share of final demand and $b_{(I)}$ is a vector whose ith element is b_i/b_I if i belongs to industry I and zero otherwise. The scalar Θ is defined in the proof of the proposition, and represents the gap between real GDP and welfare in a version of the model with only industry-level shocks.

In words, Proposition 13 implies that if firms' productivity and demand shocks are correlated with each other (but not necessarily across firms), then there is a gap between real GDP and welfare that does not appear in an industry-level specification of the model. Furthermore, this bias is, to a second-order, additive. That is, the overall bias is the sum of the industry-level bias (that we studied in the previous section) plus the additional bias driven by within-industry covariance of supply and demand shocks. Note that if supply and demand shocks at the firm level are correlated and persistent, then the bias grows over time, as in our product-level data discussed below.

Proof of Proposition 13. Start by setting nominal GDP to be the numeraire. To model the industry-structure, for each industry I, add two new CES aggregators. One buys the good for the household and one buys the good for firms. Let firm $i's$ share of industry I from household expenditures be b_{iI}. Let the expenditure share of other firms on firm i be s_{iI}. We have

$$\sum_{i \in I} b_{iI} = 1$$
$$\sum_{i \in I} s_{iI} = 1.$$

Let λ_I^c and λ_I^f be sales of industry I to households and firms. Then we have

$$d\lambda_I = d\lambda_I^c + d\lambda_I^f.$$

The sales of an individual firm i in industry I is given by

$$\lambda_i = b_{iI}\lambda_I^c + s_{iI}\lambda_I^f, \tag{55}$$

$$d\lambda_i = db_{iI}\lambda_I^c + b_{iI}d\lambda_I^c + ds_{iI}\lambda_I^f + s_{iI}d\lambda_I^f, \tag{56}$$

$$db_{iI} = Cov_{b_I}(d\log x + (1-\zeta_I)d\log A, Id_{(:,i)}),$$

$$ds_{iI} = Cov_{s_I}((1-\sigma_I)d\log A, Id_{(:,i)}),$$

where $Id_{(:,i)}$ is a vector of all zeros except for its ith element which is equal to one, b_I is a vector of market shares in final sales of industry I, and s_I is a vector of market shares in non-final sales of industry I.

The gap between macro welfare and real GDP, $EV^M - \Delta\log Y$, is approximately given by

$$\frac{1}{2}d\log x\frac{\partial\lambda}{\partial\log x}d\log A = \frac{1}{2}\sum_{i\in N}\left[\sum_{j\in N}d\log x_j\frac{\partial\lambda_i}{\partial\log x_j}\right]d\log A_i.$$

Using (56), the sums can be re-written as

$$\sum_{i\in N}\left[\sum_{j\in N}d\log x_j\frac{\partial\lambda_i}{\partial\log x_j}\right]d\log A_i = \sum_{i\in N}\left[d\log x\frac{\partial b_{iI}}{\partial\log x}\lambda_I^c d\log A_i + b_{iI}d\log x\frac{\partial\lambda_I^c}{\partial\log x}d\log A_i\right.$$
$$\left.+d\log x\frac{\partial s_{iI}}{\partial\log x}\lambda_I^f d\log A_i + s_{iI}d\log x\frac{\partial\lambda_I^f}{\partial\log x}d\log A_i\right],$$

where now the subscript I indicates the industry that the firm i belongs to.

The individual terms can be written out as

$$\begin{aligned}\sum_{i\in N}\left[d\log x\frac{\partial b_{iI}}{\partial\log x}\lambda_I^c d\log A_i\right] &= \sum_{i\in N}Cov_{b_I}(d\log x, Id_{(:,i)})\lambda_I^c d\log A_i\\ &= Cov_{b_I}(d\log x, \sum_{i\in N}Id_{(:,i)}d\log A_i)\lambda_I^c\\ &= Cov_{b_I}(d\log x, d\log A)\lambda_I^c;\end{aligned}$$

$$\sum_{i\in N}\left[b_{iI}d\log x\frac{\partial\lambda_I^c}{\partial\log x}d\log A_i\right] = \mathbb{E}_{b_I}\left(d\log A\right)d\log x\frac{\partial\lambda_I^c}{\partial\log x};$$

$$\sum_{i\in N}d\log x\frac{\partial s_{iI}}{\partial\log x}\lambda_I^f d\log A_i = 0;$$

and

$$\sum_i s_{iI} d\log x \frac{\partial \lambda_I^f}{\partial \log x} d\log A_i = \mathbb{E}_{s_I}\left(d\log A\right) d\log x \frac{\partial \lambda_I^f}{\partial \log x}.$$

Of the four terms, two depend on changes on industry-level sales shares, one of them is zero, and the remaining one (the first term) is the within-industry covariance of supply and demand shocks that is highlighted in the statement of the proposition. Hence, the remaining terms in the statement of the proposition are

$$\Theta = \sum_I \left[\mathbb{E}_{s_I}\left(d\log A\right) d\log x \frac{\partial \lambda_I^f}{\partial \log x} + \mathbb{E}_{b_I}\left(d\log A\right) d\log x \frac{\partial \lambda_I^c}{\partial \log x}\right].$$

□

Appendix J Additional Details on Nielsen Application

In this appendix, we provide additional details on how we treat the data when constructing Figure 3, and we perform some robustness exercises with respect to the elasticity of substitution.

Details on the construction of Figure 3 The Nielsen Consumer Panel data are provided under subscription through the Kilts Center for Marketing at the University of Chicago. A first file provides quantity and expenditures net of discount by UPC (universal product code) for each shopping trip recorded by roughly 60,000 households in the panel.[54] Additional files record the date of each shopping trip and describe household characteristics, including the Nielsen-defined market in which each household resides. Nielsen provides a set of weights so that each household in the panel can be understood to represent a certain number of households in their market for a given panel year. Nielsen also provides a file with descriptions of each product, including a set of Nielsen-defined product categories. The lowest level of product categorization in this scheme is known as a module. The Kilts Center tracks UPCs over time, assigning UPC version numbers that record if characteristics associated with a given barcode change over time. Thus, a UPC-version has a fixed set of product characteristics over time, and we use this stable-characteristic notion of UPCs.

This makes it unlikely that the good undergoes quality changes over time. First, as pointed out by Redding and Weinstein (2020), this is because firms prefer to use different barcodes for products with different observable characteristics for inventory and stock con-

[54] 40,000 households in panel years 2004 to 2006.

trol purposes. Second, even if a product keeps the same barcode but undergoes a change in one of the observable characteristics tracked by the Kilts Center, then it is not treated as the same product in our sample.

We construct our sample as follows. After dropping trips with non-positive quantity or non-positive expenditure net of discounts, we collapse household-trip-UPC observations by summing to household-quarter-UPC observations. For each household-quarter-UPC, we calculate the average unit value (expenditures/quantity) and drop observations that are more than three times or less than one third the median unit value for observations in the same market-quarter-UPC, as well as those for which the quantity purchased is more than 24 times the median within the same market-quarter-UPC.

In turn, we collapse the household-quarter-UPC data to a year-UPC panel by summing (scaled by the Nielsen household projection factor) quantities and expenditures by UPC and by year. Annual price is defined as the ratio of annual expenditures and annual quantity.

We calculate the growth rate of each good's price and expenditure between adjacent years (e.g. 2013 price / 2012 price), and identify observations with "extreme growth rates" as instances where the price and/or expenditure growth rate are outside the 1st and 99th percentiles among all year-to-year price and expenditure growth rates for goods with non-zero expenditures in all 8 quarters in adjacent years.

We set $t_1 = 2019$, and $t_0 = 2004, ..., 2018$. For each t_0 we construct a balanced sample of UPCs with non-extreme growth rates and non-zero expenditures in every quarter between t_0 and 2019. In addition, we impose a balanced panel of modules that have at least two unique UPCs available in every quarter from 2004 to 2019. This panel of modules also excludes so-called magnet series and "unclassified" module categories. For $t_0 = 2018$, the balanced sample includes 822 modules and $247,611$ products (average of 301 products per module, median of 137 products per module). For $t_0 = 2004$, the balanced sample includes the same 822 modules and $32,030$ products (average of 39 products per module, median of 17 products per module).

For each t_0 (x-axis in the figure) we construct chained-Tornqvist, long-difference Sato-Vartia, and "welfare-relevant" (equivalent variation at $t_1 = 2019$ preferences) prices indices for each module including only those goods in the corresponding t_0 balanced sample. These module price indices are combined into a single aggregate index by weighting each module's price index by its share of expenditure in 2019 among modules in the balanced panel (i.e. every set of t_0 module price indices is aggregated using the same weights, and weights sum to unity for the 822 modules in the sample). For the chained-Tornqvist, for each module we construct year-by-year Tornqvist price indices and cumulate them between t_0 and t_1. For the long-difference Sato-Vartia, we apply the standard formula to a single change in prices and expenditures between t_0 and t_1 (in contrast to the chained-Tornqvist, which uses prices and expenditures in every year between t_0 and t_1). For welfare, we assume for each module a homothetic-CES aggregator with elasticity of substitution $\theta_0 = 4.5$ (we report robustness to lower and higher values of θ_0). The welfare-relevant price index based on $t_1 = 2019$ preferences, given price changes between t_0 and t_1, is

$$\log\left(\sum_i b_{it_1}\left(\frac{p_{it_1}}{p_{it_0}}\right)^{1-\theta_0}\right)^{\frac{1}{1-\theta_0}}$$

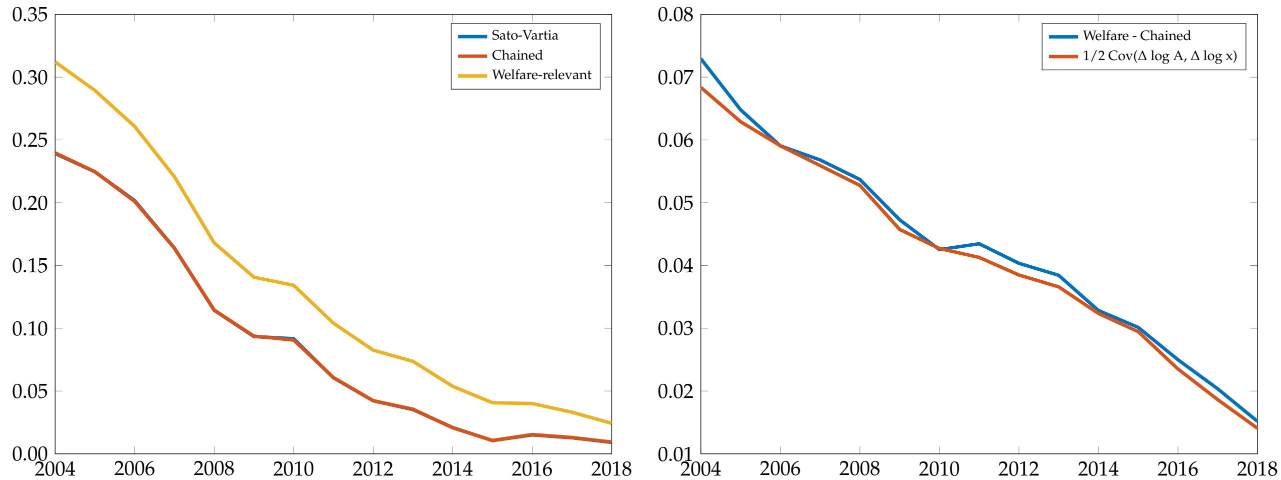
where b_{it_1} denotes the t_1 budget share of good i within its module among goods in the t_0-continuing goods sample.

The first panel of Figure 3 reports all three price indices for $t_0 = 2004, ..., 2018$. Note that, for each t_0, all three price indices are based on the same sample of products but the sample varies with t_0 due to product entry and exit.[55]

The second panel of Figure 3 compares the gap between the "welfare-relevant" series and the "chained" series to the gap implied by the approximation formula in Proposition 13. Specifically, we report the term $\frac{1}{2}\sum_I b_I Cov_{b_{(I)}}(\Delta \log x, \Delta \log A)$, again weighting modules by their spending share in 2019. We construct $\Delta \log A$ as the difference between each good's own log-price change and the module's average log-price change (for goods in the corresponding balanced sample). We construct Δx as the difference between the observed change in expenditure shares and the change in expenditure share implied by a CES aggregator with elasticity $\theta_0 = 4.5$ given observed price changes.

Robustness Figures 6 and 7 replicate Figure 3 using lower and higher values for the elasticity of substitution. The size of the bias gets smaller as we get closer to Cobb-Douglas. This is because in the data changes in prices and changes in expenditure shares are approximately uncorrelated. When demand is Cobb-Douglas, changes in expenditure shares are taste shocks, and since taste shocks are uncorrelated with price changes, following the logic of Proposition 2, the bias is small.

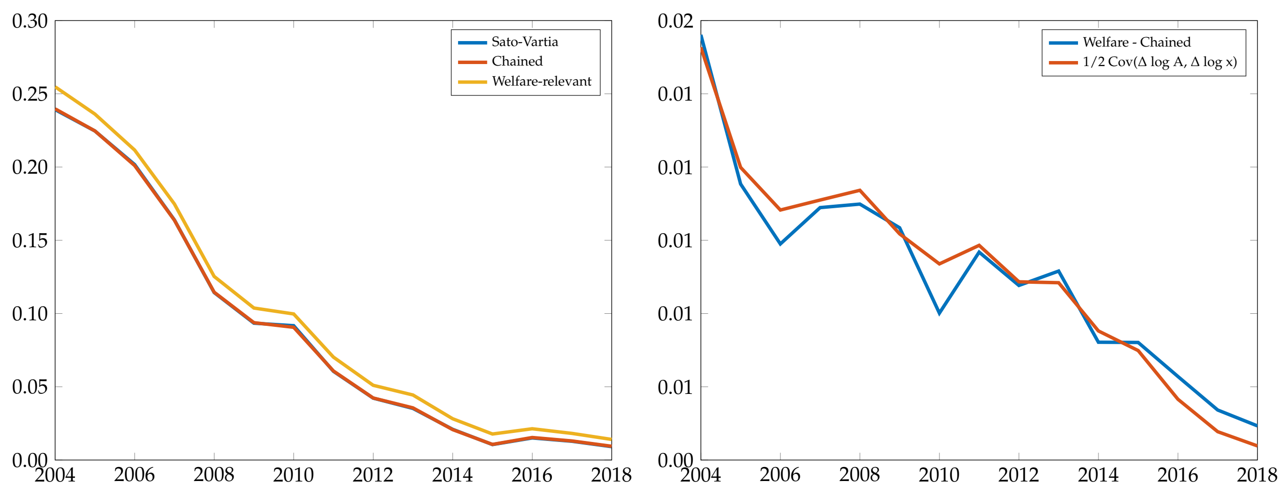
[55] We consider an alternative chained-price index which constructs year-by-year Tornqvist price indices using products with non-extreme growth rates and nonzero expenditures in all 8 quarters in these two adjacent years (but not balanced in the overall period between t_0 and t_1) which we then cumulate between t_0 and t_1. The resulting inflation is lower than using our balanced Tornqvist index, implying an even larger gap between chained and welfare-relevant inflation. We choose the balanced Tornqvist index as a baseline so that all three indices are based on the same set of observations for any given t_0, and because in our welfare measure we abstract from entry and exit.



(a) Inflation rates

(b) Bias

Figure 6: Welfare-relevant, chain-weighted, and Sato-Vartia inflation rate for continuing products. The welfare-relevant rate is computed assuming that the elasticity of substitution across UPCs in the same module is 6.5.



(a) Inflation rates

(b) Bias

Figure 7: Welfare-relevant, chain-weighted, and Sato-Vartia inflation rate for continuing products. The welfare-relevant rate is computed assuming that the elasticity of substitution across UPCs in the same module is 2.5.

Appendix K Non-CES Functional Forms

In this appendix, we generalize Proposition 7 beyond CES functional forms. To do this, for each producer k with cost function $\mathbf{C}_k$, we define the Allen-Uzawa elasticity of substitution between inputs x and y as

$$\theta_k(x,y) = \frac{\mathbf{C}_k d^2\mathbf{C}_k/(dp_x dp_y)}{(d\mathbf{C}_k/dp_x)(d\mathbf{C}_k/dp_y)} = \frac{\epsilon_k(x,y)}{\Omega_{ky}},$$

where $\epsilon_k(x,y)$ is the elasticity of the demand by producer k for input x with respect to the price p_y of input y, and Ω_{ky} is the expenditure share in cost of input y. For the household $k = 0$, we use the household's expenditure function in place of the cost function (where the Allen-Uzawa elasticities are disciplined by Hicksian cross-price elasticities and expenditure shares).

Following Baqaee and Farhi (2019b), define the *input-output substitution operator* for producer k as

$$\Phi_k(\Psi_{(i)}, \Psi_{(j)}) = -\sum_{1\leq x,y\leq N+1+F} \Omega_{kx}[\delta_{xy} + \Omega_{ky}(\theta_k(x,y) - 1)]\Psi_{xi}\Psi_{yj}, \tag{57}$$

$$\tag{58}$$

where δ_{xy} is the Kronecker delta. Then, Proposition 7 generalizes as follows:

Proposition 14. *At any point in time t, changes in the relevant variables are pinned down by the following system of equations*

$$d\log p_{it} = -\sum_{j\in N} \Psi_{ijt} d\log A_{jt} + \sum_{f\in F} \Psi^F_{ift} d\log \lambda_{ft}. \tag{59}$$

Changes in sales shares for goods and factors are

$$\begin{aligned}\lambda_{it} d\log\lambda_{it} = &\sum_{j\in\{0\}+N} \lambda_{jt}\Phi_j\left(-d\log p_t, \Psi_{(:,t),t}\right) \\ &+ Cov_{\Omega_{(0,:),t}}\left(d\log x_t, \Psi_{(:,i),t}\right) + Cov_{\Omega_{(0,:),t}}(\varepsilon_t, \Psi_{(:,i),t})\left(\sum_{k\in N}\lambda_{kt} d\log A_{kt}\right).\end{aligned} \tag{60}$$

Changes in welfare-relevant variables are pinned down by the same set of differential equations above where the second line of (60) *is set to zero and the boundary conditions are that $\Omega = \Omega_{t_1}$ and $\Psi = \Psi_{t_1}$.*

Since Φ_j shares many of the same properties as a covariance (it is bilinear and symmetric in its arguments, and is equal to zero whenever one of the arguments is a constant), the intuition for Proposition 14 is very similar to that of Proposition 7. Computing the equilibrium response in Proposition 14 requires solving a linear system exactly as in Proposition 7.

Appendix L Distorted Economies

In this appendix, we show how to extend the results in Section 3 to economies with inefficient equilibria building on the results of Baqaee and Farhi (2019a) for economies with homothetic and stable preferences. Consider again the environment in Section 3, but suppose that there are some arbitrary pattern of distorting wedges at point μ, which are implicit or explicit taxes. Without loss of generality, we can assume that μ take the form of output wedges (i.e. a tax wedge between price and marginal cost).[56]

For each A, x and μ, we denote equilibrium prices and aggregate income by $p(A, x, \mu)$ and $I(A, x, \mu)$. These equilibrium prices and incomes are unique up to the choice of a numeraire. Define the *macro indirect utility* function $V(A, L, \mu; x)$ to be the utility achieved by the agent with preferences x under the Walrasian equilibrium with wedges.

Consider shifts in technologies from A_{t_0} to A_{t_1}, along with changes in preferences from x_{t_0} to x_{t_1} and output wedges from μ_{t_0} to μ_{t_1}. We use the same definition of welfare as in Section 3, but we no longer require that the first welfare theorem hold.

We now characterize changes in real GDP and welfare. For simplicity, we abstract from changes in factor quantities, L. As in Section 2, to study this problem we index the path of technologies, preferences, and wedges by time t. The definition of $\Delta \log Y$ is the same as before: $\Delta \log Y = \int_{t_0}^{t_1} \sum_{i \in N} b_{it} d \log c_{it}$. Define $\tilde{\lambda}$ to be the cost-based Domar weight of i, as in Baqaee and Farhi (2019a). That is,

$$\tilde{\lambda}' = b'(I - \mu\Omega)^{-1},$$

where b, μ, and Ω are all functions of A, u, x, and μ.

Proposition 15 (Real GDP)**.** *Given a path of technologies, tastes, and wedges that unfold as a function of time t, the change in real GDP is*

$$\Delta \log Y = \int_{t_0}^{t_1} \sum_{i \in N} \tilde{\lambda}_i(A_t, x_t, \mu_t) \frac{d \log A_{it}/\mu_{it}}{dt} - \int_{t_0}^{t_1} \sum_{i \in F} \tilde{\lambda}_i(A_t, x_t, \mu_t) \frac{d \log \lambda_{it}}{dt} dt. \tag{61}$$

Define $\lambda^{ev}(A, \mu)$ to be sales shares in a fictional economy with productivities A and wedges μ, but where consumers have stable homothetic preferences represented by the expenditure function $e^{ev}(p, u) = e(p, u_{t_1}, x_{t_1}) \frac{u}{u_{t_1}}$ where $u_{t_1} = v(p_{t_1}, I_{t_1}; x_{t_1})$, similar to Section 2. Let $\tilde{\lambda}^{ev}$ be the equivalent cost-based Domar weights.

[56] This is without loss of generality because we can always introduce a wedge on i's purchases of inputs from j by adding a fictitious middle-man that buys from j on behalf of i. An output wedge on this fictitious middleman is isomorphic to an input-specific wedge in the original economy.

Proposition 16 (Macro Welfare)**.** *For any smooth path of technologies, tastes, and wedges that unfold as a function of time t, changes in macro welfare are*

$$EV^M = \int_{t_0}^{t_1} \sum_{i \in N} \tilde{\lambda}_i^{ev}(A_t, \mu_t) \frac{d \log A_{it}/\mu_{it}}{dt} dt - \int_{t_0}^{t_1} \sum_{i \in F} \tilde{\lambda}_i^{ev}(A_t, \mu_t) \frac{d \log \lambda_{it}^{ev}}{dt} dt. \tag{62}$$